Set Name side by side	Query	Hit Count	Set Name result set
DB=USPT; PLUR=YES; OP=ADJ			
L6	L5 and (address$ with extract$)	1	L6
L5	L4 and (compar$ or match$)	12	L5
L4	L3 and (junk or spam).ab.	14	L4
L3	L2 and ((email$) or (e-mail$)).ab.	275	L3
L2	((709/$)!.CCLS.)	15684	L2
L1	709/$/ccls.	0	L1

END OF SEARCH HISTORY

WEST

Generate Collection | Print

Search Results - Record(s) 1 through 10 of 12 returned.

☐ 1. Document ID: US 6546416 B1

L5: Entry 1 of 12 File: USPT Apr 8, 2003

DOCUMENT-IDENTIFIER: US 6546416 B1
TITLE: Method and system for selectively blocking delivery of bulk electronic mail

Abstract Text (1):
The origin address of an e-mail message is validated to enable blocking of e-mail from spam e-mail sources, by preparing, in response to the receipt of a predetermined e-mail message from an unverified source address, a data key encoding information reflective of the predetermined e-mail message. This message, including the data key, is then issued to the unverified source address. The computer system then operates to detect whether a response e-mail message, responsive to the challenge e-mail message, is received and whether the response e-mail message includes a response key encoding predetermined information reflective of a predetermined aspect of the challenge e-mail message. The unverified source address may be recorded in a verified source address list. Thus, when an e-mail message is received, the computer may operate to accept receipt of a predetermined e-mail message on condition that the source address of the predetermined e-mail message is recorded in the verified source address list and alternatively on condition that the predetermined e-mail message includes the response key.

Detailed Description Text (15):
The accept and reject lists 24, 26 provide storage for respective lists of e-mail addresses preferably on the local e-mail client computer system 14. The form of the addresses as stored may include simple domain names, specific user e-mail addresses, and Internet protocol (IP) numbers. Inclusion and exclusion operators, wildcards and IP range lists may also be utilized in the parsing or other evaluation of the accept and reject lists address. The use of such operators, wild cards and lists in considering whether a specific e-mail address matches an entry in a list of e-mail addresses is known. Thus, conventional evaluation of whether a particular e-mail address matches an entry on either the accept list 24 or the reject list 26 is utilized by the present invention.

Detailed Description Text (16):
The challenge list 28 may provide storage for destination e-mail addresses of challenge messages sent (not required), identifiers of the temporarily stored messages that are being challenged (can be input from the construction of the challenge list 28' as part of the pending box 36), and certain additional information pertaining to the individual challenge messages, such as the signature encoding key and cognitive response expected for each challenge message (may alternately be determined algorithmically upon evaluation of the challenge reply message). The use of operators, wildcards, or lists are preferably not necessary in specifying e-mail address entries on the challenge list 28. Since the list 28 operates as a temporary store of information concerning the currently outstanding challenges issued by the system 22, the matching of e-mail addresses by the e-mail client system 22 against the entries in the challenge list 28 will preferably be on an exact basis.

Detailed Description Text (21):
E-mail messages not recognized as challenge reply messages at step 42, and all messages if the step 42 is not used, are then considered at step 44 to determine

whether the From or Reply-to address is present on the accept list 24. Where a match is identified, the e-mail message present in the pending box 36 is passed on to the inbox 30 for subsequent conventional processing.

Detailed Description Text (22):
If the accept list match fails at step 44, a reject step 46 is invoked to determine whether an address match can be found against the reject list 26. If a reject list match is found, the corresponding e-mail message in the pending box 36 is discarded or, in a preferred embodiment of the present invention, passed to the discard box 32 for subsequent conventional processing.

Detailed Description Text (23):
If a reject list match is not found, the message content is preferably evaluated partially through the step 48 to determine whether, for example, a known correspondent is replying to an e-mail message originated from the system 22', but replied to from an e-mail account not previously seen by the system 22'. Since out-bound messages from the e-mail user of the system 22 are preferably provided with digital signatures, responses to such messages are validated and thus are shown to the user when they are received. As before, the digital signature preferably encodes the date that the message was sent. Thus, the step 48 can be set to invalidate messages received beyond a nominal reply period determinable by the e-mail user of the system 22. Preferably, messages containing expired digital signatures are discarded or put in the trash box 32; validated messages are passed to the inbox 30. In a preferred embodiment of the present invention, the address of e-mail messages validated only by virtue of a valid digital signature are not placed on the accept list. Rather, no present action is taken regarding messages from this address, thereby permitting the active challenge system 22' to re-evaluate messages received subsequently from that address. The accept list will be updated with this address if, however, the e-mail user chooses to update the list 24 or the e-mail user simply replies directly to this address.

Detailed Description Text (24):
Finally, messages received but not matched to the accept or reject lists and not containing a digital signature are, in a step 50, responded to by the preparation and issuance of a challenge message. This message, once generated to include a cognitive request and a current corresponding digital signature, is placed in the out box 34.

Detailed Description Text (32):
Received e-mail messages that bear a signature but fail in the validation of the signature or are received late relative to the time threshold established directly or indirectly by the end user is identified as invalid signed message 82. The invalid received e-mail message is then discarded 76 from the temporary queue 36. Optionally, such invalid signed messages may be further evaluated to identify the sender e-mail address, which may then be added to the reject list 26. Preferably, this option is established directly or indirectly by the end user of the system 60. Conversely, where a reject list match is not found, the corresponding received e-mail message are further processed at a step 80.

Detailed Description Text (38):
A preferred process of handling original outbound messages in accordance with the preferred embodiments of the present invention is shown in FIG. 5. The process 120 is initiated when a message is prepared 122 by the e-mail client 22'. When the message is prepared to be sent 124 by transfer 134 to the output queue 34 of the e-mail client 22, the message headers are first examined to determine whether the message qualifies as an original message. Messages identified as challenge messages are not considered original messages. Rather, new messages prepared by the e-mail user of the system 14, and ordinary reply and forward massages are considered original. The destination e-mail address specified in an original message is then matched 126 against the accept list 24 to determine whether the address has already been recorded. If not, the recipient e-mail address is added 128 to the accept list 24. This ensures that e-mail destinations implicitly recognized and validated by the user of the e-mail client system 22' are subsequently recognized as valid senders of e-mail messages to the system 22. In either event, a new digital signature is prepared 130 and appended 132 to the outbound message. Transfer of the resulting

message to the out-box 134 is then complete. The outbound message, along with any other pending outbound messages are subsequently picked up or transferred 136 to the ISP servicing the e-mail client 22.

Detailed Description Text (43):
Referring again to FIG. 3, in the ongoing operation of the system 60 the majority of received e-mail messages will likely be transferred 70 to the input queue 30 based on e-mail address matches against the accept list 24. In accordance with a preferred embodiment of the present invention, a quick initial development of the accept list 24 can be obtained by effective assimilation of any e-mail archives kept by the user of the system 60. Presumptively, archived e-mail messages are from or are replies to valid and acceptable e-mail correspondents.

Current US Original Classification (1):
709/206

Current US Cross Reference Classification (1):
709/219

Current US Cross Reference Classification (2):
709/225

CLAIMS:

23. The system of claim 22 wherein said predetermined signature further encodes a first date and wherein said validating means compares said first date with a second date in determining whether the predetermined response message is valid.

☐ 2. Document ID: US 6453327 B1

L5: Entry 2 of 12 File: USPT Sep 17, 2002

DOCUMENT-IDENTIFIER: US 6453327 B1
TITLE: Method and apparatus for identifying and discarding junk electronic mail

Abstract Text (1):
Apparatus, methods, systems and computer program products are disclosed to provide electronic mail systems with the capability for a group of trusted users to collectively determine whether a given electronic mail message is junk e-mail. Further, if the given electronic mail message is determined to be junk mail, the e-mail systems of other trusted users in the group dispose of unviewed copies of the junk e-mail. Thus, the invention reduces the exposure of junk e-mail messages to the group of trusted users.

Drawing Description Text (10):
FIG. 9 illustrates the procedure used to compare the characteristics of e-mail messages in accordance with a preferred embodiment;

Detailed Description Text (15):
A procedure is a self-consistent sequence of steps leading to a desired result. These steps are those requiring physical manipulation of physical quantities. Usually these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. These signals are referred to as bits, values, elements, symbols characters, terms, numbers, or the like. It will be understood by those skilled in the art that all of

these and similar terms are associated with the appropriate physical quantities and are merely convenient labels applied to these quantities.

Detailed Description Text (16):
The manipulations performed by a computer in executing opcodes are often referred to in terms, such as adding or comparing, that are commonly associated with mental operations performed by a human operator. No such capability of a human operator is necessary in any of the operations described herein that form part of the present invention; the operations are machine operations. Useful machines for performing the operations of the invention include programmed general purpose digital computers or similar devices. In all cases the method of computation is distinguished from the method of operation in operating a computer. The present invention relates to method steps for operating a computer in processing electrical or other (e.g., mechanical, chemical) physical signals to generate other desired physical signals.

Detailed Description Text (42):
FIG. 9 illustrates the process used to compare two sets of identifying characteristics or a set of characteristics and an e-mail message. If the comparison is between a set of characteristics and an e-mail message, the process first determines the characteristics of the e-mail message and then performs the same comparisons as between two sets of identifying characteristics. The comparison process for two sets of identifying characteristics starts at the terminal labeled as 901. The first test 903 is whether the Message-ID: strings of both sets of characteristics are the same and not blank. If this condition is true there is a match and the process completes through the terminal labeled as 911. If this condition is false, the second test 905 is whether both Message-ID: fields are empty; the originator and subject data are the same and 80% or more of the words in the first five strings of the body text of each message (these strings are included in the identifying characteristic sets) are the same. Finally, if the characteristics do not match the process terminates through the terminal labeled as 907. Table 1 is a pseudo-code representation of the process illustrated in FIG. 9. One skilled in the art will understand that this pseudo-code is not compilable program code but a textual representation of actual compilable code used to clearly demonstrate the operation of the process illustrated in FIG. 9. The pseudo-code shown in Table 1 is self-documenting to one skilled in the art so long as it is understood that the compare5lines function returns an integer representing the percentage of identical words within the text lines of the two sets of identifying characteristics.

Detailed Description Text (43):
Having described the databases, their maintenance processes and the process used for comparing e-mail identifying characteristics, we now discuss additional aspects of the invention within the receiving e-mail systems.

Detailed Description Text (53):
If 1105 the received message is a Junk Mail Report message, the process then checks 1111 whether the originator of the received message is a member of the trusted group by comparing the field-body text of the Sender: header field with the records 510 in the Server's Trusted Group database to determine whether the originator's e-mail address is contained in the "E-mail Address of Trusted Group Member" field 511 of one of the records 510 in the database--if the received message does not include a Sender: field, the information i n the field-body of the From: header is used. If no record 510 exists matching the originator's e-mail address, the received message is discarded 1117 and the process competes through the terminal labeled as 1109.

Detailed Description Text (55):
If 1113 the message is authentic, the process continues through the terminal labeled as 1115 to the terminal labeled as 1131 of FIG. 11b. Next 1133, the process extracts the identifying characteristics of the reported junk e-mail message and determines 1135 whether the Server's Junk E-mail database contains a record 610 with matching characteristics in the "Junk E-mail Characteristics" field 611. At this point, the received message is deleted after retaining any needed information. If 1135 such a record 610 exists, the value in the "Number of Trusted Group Reporting" field 613 is incremented 1137 and the current date is stored in the "Date of Last Report" field 615. Storing the current date in the record 610 resets this record's trip time as

used in the process described for FIG. 8. The process continues to the decision block labeled as 1141 and described below. If 1135 no record 610 matches the report's characteristics, the process creates 1139 a new record 610 in the Server's Junk E-mail database storing the characteristics of the message in the "Junk E-mail Characteristics" field 611, setting the "Number of Trusted Group Reporting" field 613 to a value of "1" and setting the "Date of Last Report" field 615 to the current date. Next 1141, the process checks whether the value contained in the "Number of Trusted Group Reporting" field 613 is greater than a preference value--if "Number of Trusted Group Reporting" field 613 is not greater than the preference value, the process completes through the terminal labeled as 1145.

Detailed Description Text (61):
However, if 1205 the new message is not a Junk Mail Warning message, the process then determines 1219 the message's characteristics. Once the new message's characteristics are obtained the process checks whether 1221 a record 600 exists in the User's Junk E-mail database that matches the new message's characteristics--if 1221 no match exists, the process continues with normal e-mail processing through the terminal labeled as 1217 and inventive aspects of the process complete.

Detailed Description Text (62):
However, if 1221 the characteristics of the new message do match the "Junk E-mail Characteristics" field 601 of a record 600, the new message has been determined to be a junk e-mail message and the new message is deleted 1223 and is thus prevented from being presented to the recipient. Thus, the invention has intercepted and disposed of the junk e-mail. Finally, the matching record 600 is updated 1225 by storing the current date in the "Last Date" field 603 of the matching record 600 and the process completes through the terminal labeled as 1213.

Detailed Description Text (64):
Finally, the new Junk Mail Warning message is discarded 1257 and the process completes through the terminal labeled as 1259. If 1253 no record 600 exists in the User's Junk E-mail database, the process creates 1261 such a record 600 and initializes its fields. The "Junk E-mail Characteristics" field 601 is initialized with the set of characteristics provided in the field-body portion of the X-Junk-Mail-Warning: header. The "Last Date" field 603 is initialized with the current date. Finally, the inbox is scanned. This scan process examines each unread e-mail message in the user's inbox, characterizes the unread e-mail message, and if the characterization matches that included in the new Junk Mail Warning message the unviewed message is deleted 1263 from the inbox so that it will not be presented to the recipient, the Junk Mail Warning message is discarded 1257, and the process completes through the terminal labeled as 1259.

Detailed Description Text (67):
This signature is a large binary number that is converted to an ASCII representation as shown in Table 3. The recipient first converts the ASCII representation back to a number and decrypts the message digest using the senders public key. Then the recipient creates another message digest of the originator's message and compares the newly created digest with the decrypted version. If the two digests are the same the message is authentic.

Detailed Description Paragraph Table (1):
TABLE 1 MessageID_type: messageIDA, messageIDB; boolean match = FALSE; if ((messageIDA not empty) and (messageIDB not empty)) { if (messageIDA equal messageIDB) {match = TRUE}; } elseif ((messageIDA empty) and (messageIDB empty)) { if ((messageIDA.send equal messageIDB.sender) and (messageIDA.subject equal messageIDB.subject)) { if (80 lessthanorequal compare5lines(messageIDA.text, messageIDB.text)) {match = TRUE}; } }

Current US Cross Reference Classification (2):
709/206

☐ 3. Document ID: US 6421709 B1

L5: Entry 3 of 12 File: USPT Jul 16, 2002

DOCUMENT-IDENTIFIER: US 6421709 B1
TITLE: E-mail filter and method thereof

Abstract Text (1):
A system and method of filtering junk e-mails. A user is provided with or compiles a list of e-mail addresses or character strings which a user would not wish to receive to produce a first filter. A second filter is provided including names and character strings which the user wishes to receive. Any e-mail addresses or strings contained in the first filter will be automatically eliminated from the user's system. Any e-mail addresses or strings contained in the second filter would be automatically sent to the user's "in box". Any e-mail not provided in either of the filtered lists will be sent to a "waiting room" for user review. This user review results in the user rejecting any e-mail, the addresses as well as specific character strings included in this e-mail would be transmitted to a central location to be included in a master list. This master list is periodically sent to each of the users allowing the first filter to be updated. A collaborative filter is used employing message base filtering that is not effected by e-mail header forgery and utilizes the networked intelligence of end users to maintain a highly inaccurate and comprehensive filter. The collaborative filter would then use the real-time input from the end users to keep the users involved in the filtering process.

Detailed Description Text (3):
The software would allow the individual user to construct an automatic discard filter 12. The automatic discard filter is a collective term consisting of a user modified discard filter, a user personal address filter as well as a user personal string filter. During operation of the system, the automatic discard filter 12 would include a current filter list comprising a list of active e-mail addresses against which incoming e-mails are compared. This current filter list is retained in a memory section of the users computer. Any comparison between any incoming e-mail and the current filter list could be accomplished within the user's computer system. The current filter list is maintained at the remote central location 46 as well as being periodically updated in each of the users PC systems 48. The remote location 46 would include a delta filter server 22 and download server 24 for a particular user as well as delta server filter 26 from all other users. The current filter list can be modified by the user to personally remove any addresses therefrom through various deletion techniques, thereby providing the user with a user modified discard filter. The user personal address filter would include additional addresses the user has added to the current filter list as well as any character strings that the user has added via a text entry containing an "@". For the purpose of the present invention, a text entry is a character string entered into the system by keyboard typing. Typing is initiated by double clicking or highlighting and typing, thereby clearing an old string and creating a new string. When the mouse is clicked on some other location or "enter" is hit, the string will be entered into the appropriate memory structure for the new field.

Detailed Description Text (6):
Any e-mail received by the user is checked against the automatic discard filter 12 to determine whether any character string on the "No Admittance List" 52 will bar entry of any e-mail with matching text in its address, subject line or message body. If this occurs, that e-mail will be eliminated from the users system as indicated by the Trash Bin 16.

Detailed Description Text (7):
Conversely, any address contained in the Guest List which matches an address of an incoming e-mail would be automatically forwarded to the In Box folder 18 for review by the user. Similar to the situation with respect to the No Admittance List 52, a text string entered in the Guest List 54 would forward all messages containing that character string to the "In Box" folder 18. This feature would allow users to

receive on-demand direct marketing information from parties, promoting products for which the user has expressed interest based upon the text string entered in the Guest List 54.

Detailed Description Text (10):
Periodically, the database server 24 in communication with the address filter server 22 would download updated filter addresses to the various users in the system by constructing an address packet consisting of every address on the current filter list since the date and time of each of the users last update. The address packet is a data structure consisting of N strings of e-mail addresses and a variable containing the time of construction of the packet. The packet is compressed for downloading and uploading multiple e-mail addresses. Based upon the particular implementation of the software of the present invention, the updated version of the current filter list is substituted for the No Admittance List currently provided in the users system. Alternatively, since the No Admittance List might include addresses and character strings personally added by the user but not included in the current filter list, the updated filter list would be compared with the automatic discard filter and any additional entries not included in the automatic discard filter would be added thereto.

Detailed Description Text (11):
FIG. 2 illustrates the In Box folder 18 and the Waiting Room 20 in more detail as well as giving examples of the type of messages included therein. The list of names included ES6 on the automatic discard filter 12 are provided in the No Admittance List 20. Any incoming e-mail whose new address matches one of the addresses on this list is immediately discarded to Trash 16. Addresses may be added to this list via an update button 61, the Add to No Admittance button 58, text entry, or by dragging a selected e-mail to this window with the mouse. The update button 62 automatically downloads the latest automatic discard filter from the download server 24. The updated filter list is displayed in the No Admittance Window. Simultaneously, user added e-mail addresses are sent to the Delta Filter Server 22 for consideration in future updates to the users in the system.

Detailed Description Text (15):
The Guest List window 54 would include a list of names on the Guest List filter. Any incoming e-mail whose new address matches one of the addresses on this list is immediately forwarded to the In Box folder 18. Addresses may be added to this list via the add to Guest List button 60, text entry, or by dragging a selected e-mail to this window with a mouse.

Detailed Description Text (16):
The In Box 18 includes only those e-mails that have successfully passed through both the automatic discard filter and the Guest List Filter. Additionally, any e-mail from any folder, may be selected and dragged into the In Box 18 by the user using the mouse. Similar to the Waiting Room 20 the In Box 18 includes the e-mail addresses, the date and time of receipt as well as the subject matter of the e-mail. Furthermore, the particular configuration of the In Box 18 as illustrated in FIG. 2 can be changed depending upon the users requirements. By clicking on an open slot in the No Admittance List 52 or the Guest List 54 or by double clicking on a existing text, the user may enter a character string to be checked in the filtering system. Any such character string on the No Admittance List will bar the entry of any e-mail with matching text in its address, subject line or message body. For example, as shown in the No Admittance List 52, any received e-mail with the words "free money" in its subject or message would be discarded. A text string similarly entered in the Guest List would forward all messages containing that character string to the In Box. Text entry can also be used to type in new e-mail addresses or edit existing ones on either of the filter lists.

Detailed Description Text (22):
Both the control screen for the member database 62 and the address database 70 contain a search field 64, a search panel 66 and a filter panel 68. The search field 64 would contain information matching an entry in either the address database or the member database. Buttons 72 and 74 would allow either of these databases to appear on the control screen. Both of these databases would include search results run in either the address database or the member database in Section 66. The current filter

Section 68 would allow entries to be updated or saved at various times. It would also include a box 76 indicating the number of days an address can remain on the current filter list without a new instance of that address being uploaded by the filter users. It would also include a box 78 listing the minimum number of reportings required for an address to be placed on the current filter list. Certainly both of these central screens can be set up in multitude of ways depending upon the specific information to be provided.

Detailed Description Text (23):
Returning to FIG. 1, the process of comparing received e-mails to both the Automatic Discard filter 12 and the Guest List filter 14 will now be explained. Incoming e-mails 28, 30 and 32 are compared to information contained in the user modified discard filter, the user personal address filter and the users personal string filter utilizing the address line, the subject line as well as the message body. Since the information included in e-mails 28, 30 and 32 are not contained in the automatic discard filter, all three of these e-mails are directly transmitted to the Guest List filter 14. The e-mail addresses, subject line and message body of these three e-mails result in a match for all three of these e-mails. Consequently, these e-mails are sent to the In Box folder 18.

Detailed Description Text (33):
E-mail is then distributed to the end user mailboxes. When users download their e-mail from the Mail Drop Service, it is filtered again via the Spam Filter. Those that match are discarded before the users ever see them. All other messages pass safely through to the user. It is necessary to filter at both the Mail Server and Mail Drop Service level to ensure that end users are protected by the latest updates to the Collaborative Filter. However, the system would still operate if no filtering were to be done at the Mail Drop Service level.

Detailed Description Text (37):
Initially, the spam message enters the system. User A logs on first and checks his e-mail. User A downloads two messages: his non-spam message and a spam message. Note that the spam message passes through the filter since it has never been seen before. User A notes the spam message and submits it to the Collaborative Filter via a simple button click. The Collaborative Filter uses User A's submission to update the filter. From this point on, any new incoming message that matches the submitted spam e-mail will be discarded.

Detailed Description Text (47):
The spamCheck( ) function checks whether a given e-mail is junk in the following fashion: 1. spamCheck( ) generates a signature for the message. 2. spamCheck( ) queries the Spam Database for the message's signature. 3. If the database query does not find the message's signature, then the e-mail is not junk and it an be passed on to end users. 4. If the database query does find the message's signature, then a matching function is used to determine whether or not the message in questions truly matches a message in the Spam Filter database. a. If the matching function does not find these messages to be equivalent, then the message is not junk. b. If the matching function does find these messages to be equivalent, then the message is filtered.

Detailed Description Text (49):
The signature described in the above process is a hash function based on the message's body. Message signatures are very important because they allow the Collaborative Filter to operate in an efficient manner. The message signature enables non-junk e-mail messages to quickly pass through the filter. This occurs because it is extremely unlikely that an incoming non-junk message signature will match the signature of a junk e-mail already stored in the database.

Detailed Description Text (50):
Since the message signature is a type of hashing function, there will be some unavoidable signature collisions (i.e., two unique messages which generate the same signature). The filtering algorithm resolves signature collisions by calculating a matching function on both messages to ascertain if these messages are really equal. The matching function uses a combination of techniques (e.g., checksum, fuzzy matching) to generate a likelihood that two messages are essentially equivalent.

Exact comparisons cannot be used since junk e-mail senders will embed extra characters in their outgoing e-mails to circumvent message based filtering techniques. For example, spammers may add extra characters at the beginning of a message by including personalized salutations. A fuzzy matching function is the appropriate solution to this problem because a spammer cannot change that portion of an e-mail's body that contains his or her message (e.g., advertisement).

Detailed Description Text (65):
A relational database system must be allocated to hold the Collaborative Filter. The system must be able to handle the extra bandwidth generated by queries form the Spam Filter library. For large organizations, it is recommended to use a dedicated database system for the Collaborative Filter. Customers should note that the Collaborative Filter has an unusual query load compared to most on-line transaction processing systems, since over 90% of its requests will be pure queries (i.e., there will be very few inserts, updates or deletes) . Due to this unique query load, a dedicated database system that can be optimized for filtering is recommended.

Current US Original Classification (1):
709/206

Current US Cross Reference Classification (1):
709/207

CLAIMS:

1. A method for mail server side filtering electronic mail received over a communication medium comprising the steps of: providing a first filter at the mail server including a list of spam messages which should not be sent to an end user; receiving a first electronic message ultimately intended for one or more end users at the mail server; comparing at the said mail server said received first electronic message to said spam messages provided in said first filter, said received first electronic message discarded if said received first electronic message is included as a spam message in said first filter and transmitting said first electronic message to the end user if said received first electronic message is not included as a spam message in said first filter; providing a second filter at the mail server for the receipt of a second electronic message sent over the communications medium from an end user, said second electronic message received at said second filter considered to be a spam message by the end user; providing a counter associated with said second filter; counting the number of repeated second electronic messages received by said second filter; adding said second electronic message to said first filter as a spam message, if said counter exceeds a predetermined value; and adding said second electronic message to said first filter as a spam message, if said second filter determines that said second electronic message is a spam message.

2. A system for filtering electronic mail received over a communication medium to an end user's computer, comprising: a mail server for receiving first electronic messages intended to be received by the end user, said mail server provided at a location remote from the end user; a first filter located at said mail server, said filter including a list of spam messages which should not be sent to the end user, said first filter also including a device for comparing at least a portion of the body of said first electronic messages with each of said spam messages; a second filter located at said mail server for receiving second electronic messages transmitted over the communication medium to said mail server from the end users; a counter and comparison device located at said mail server and in communication with said second filter for counting and classifying the number of said second electronic messages received by said second filter, said counting and comparison device determining the number of similar second electronic messages received by said second filter; wherein when the number of similar second electronic messages received by said second filter exceeds a predetermined number, said similar second electronic message is added to said first filter as an additional spam message.

☐ 4. Document ID: US 6393465 B2

L5: Entry 4 of 12 File: USPT May 21, 2002

DOCUMENT-IDENTIFIER: US 6393465 B2
TITLE: Junk electronic mail detector and eliminator

Abstract Text (1):
A method and system for parsing and analyzing incoming electronic mail messages to determine a confidence factor indicative of whether or not the messages are junk e-mail. The method and system utilize message services which attempt to contact the purported sender in order to verify that the identified host computer actually exists and accepts outgoing mail services for the specified user. The routing history is also examined to ensure that identified intermediate sites are also valid. Likewise, seed addresses can alert an e-mail provider to potential mass mailings by reporting when mail is received for ghost or non-existent accounts.

Brief Summary Text (6):
Documents are available which describe electronic mail handling procedures. In particular, two Internet standards on e-mail are incorporated herein by reference in their entirety. They are: Internet STD0014 entitled "MAIL ROUTING AND THE DOMAIN SYSTEM" (also known as RFC 974) and Internet STD0010 entitled "SIMPLE MAIL TRANSFER PROTOCOL" (also known as RFC 821). The contents of the Second Edition of "sendmail" by Bryan Costales and Eric Allman, published by O'Reilly Publishing, is also incorporated herein by reference. Further, some issued patents address the general handling of electronic mail. For example, U.S. Pat. No. 5,377,354 teaches a method for prioritizing a plurality of incoming electronic mail messages by comparing the messages with a list of key words. U.S. Pat. No. 5,619,648 teaches a method for reducing junk e-mail which uses non-address information and uses a filtering system that has access to models of the user's correspondents. The e-mail system adds a recipient identifier that is used to further specify the recipients in the group to whom the message is sent who should actually receive the message.

Detailed Description Text (4):
The method and system of the present invention assign confidence ratings to messages to signify the statuses of the messages as junk e-mails or as a bona fide messages that the recipient may wish to read. The method and system begin by analyzing the origins and transmission paths of the messages. The sender's origination information is extracted from the e-mail message and an automatic reply (called a verification request) is created and sent. Based on the verification response that is received in response to the verification request, the sender is scored as to the probable characteristics, origination, validity, and desirability of the mail. Incoming messages (e-mails) are automatically scanned and parsed, either (1) at a server located at an Internet provider (prior to delivery to the intended ultimate recipient), (2) at a LAN-based receiving station, or (3) at the actual ultimate recipient's mail machine, i.e., local to the user. Once the message has been parsed or broken down into fields, the message is compared with several user defined rules for handling messages, and a confidence rating is assigned to the message. In one embodiment, the message header information is analyzed and a verification request(s) is/are automatically sent to the purported sender(s), as identified by fields such as "From:" or "Reply-To:". If there is a delivery problem in delivering the verification request, the presumed validity of the message is reduced in accordance with a set of user-definable criteria. In addition to determining the purported origination point, the present invention automatically analyzes all information pertaining to the sender, the path of delivery, any information pertaining to copies, blind copies, or other indicia of validity of the origin of the message to determine if there has been a discernable effort to obscure the origin, disguise the sender, or in some other way to inhibit the recipient from performing verification of the sender's identity. For example, if a message has purportedly been relayed through a machine named mail.fromnowhere.com and the mail handling system has determined that such a machine does not actually exist, the confidence rating for the message should be increased.

Detailed Description Text (15):
With reference to FIG. 8, an additional level of verification is provided in the system of the present invention by including an authentication server 700 and authentication database 710. In the system of the present invention, the verification is implemented using the authenticator by calling the authentication server 700 and verifying that the id issued to the mail originator remains valid and has not been reported as a spam abuser. If a recipient has received spam, the recipient can simply put the term "spam" in the subject of the e-mail message and send it to the authentication server 700, and the id is matched with a spam designation based upon the number of mail recipients reporting the problem. After a set number (n) of spam reports, the id is assigned to the SPAM section of the authentication database 710.

Detailed Description Text (17):
In order to provide each user with an authentication ID that the authenticator can use to quickly determine if the sender is a known junk e-mailer, the e-mail users would each register, potentially for a fee, and their e-mail program would be assigned a unique identification code. Preferably, the e-mail program would maintain the unique code in secret by the mail program such that the users and others would not see the message. For example, to prevent a recipient from stealing a unique code of another user from which he/she has received a message, the e-mail program or the e-mail handling system at an ISP or corporate level could strip the unique code before delivering the message. That is, when a message is received, the mail program or mail handling system would send the unique code and the "From:" identifier to the authenticator for authentication. The code and the "From:" identifier would be checked against the database of known junk e-mailers as well as checked for consistency between the two parts. If the code was for a known junk e-mailer, or if the code and the "From:" field did not match, the mail program or mail handling system would be warned of the problem. Since the message would then be authenticated, the unique code would no longer be needed and could be stripped before passing the mail message to the user.

Current US Original Classification (1):
709/207

Current US Cross Reference Classification (1):
709/206

☐ 5. Document ID: US 6330590 B1

L5: Entry 5 of 12 File: USPT Dec 11, 2001

DOCUMENT-IDENTIFIER: US 6330590 B1
TITLE: Preventing delivery of unwanted bulk e-mail

Abstract Text (1):
Unwanted e-mail messages from bulk advertisers (SPAM) are detected and removed from a stream of e-mail, either at a central server location or an individual recipient's location. The basic on-line e-mail message, after elimination of source and addressee identification, is scanned and coded to provide a signature ID code. A set of typically three identical messages going to different e-mail addresses is detected to signify SPAM in the e-mail flow stream. Then the SPAM signature ID code is stored for use in eliminating future such messages at either a central server or one at an individual recipient's site. The signature code is typically calculated numerically, i.e. as the well known checksum in a 16-bit cyclic redundancy check routine.

Brief Summary Text (5):
Because of extremely low cost of bulk e-mail compared with other marketing strategies, such mail is not usually narrowly focussed to special audiences and results in large numbers of e-mail recipients getting messages in which they have no interest. Bulk e-mail companies thus tend to build and release mailing lists of e-mail addresses and use those lists to send messages with little discrimination or protection of the recipient's rights, desires or needs.

Brief Summary Text (17):
After detection, a numerical signature identification code for that bulk message is established, preferably by calculating a checksum using a 16-bit cyclic redundancy check. That kind of numerical signature code is quickly and easily used to compare the many messages in the e-mail flow stream without significant impedance to the flow of data.

Brief Summary Text (19):
A management coordination interface system sets up and administers the e-mail deletion process, apparatus and system. Thus a digital signature detector operating with a designated flow of e-mail sends each signature to a storage data-base comparator, which checks for uniqueness of the signature and for an initial data-base entry creates an unwanted e-mail (SPAM) signal to send to the subscribing accounts for local storage used for rejecting SPAM at authorized facilities.

Detailed Description Text (3):
As may be seen from the embodiment of FIG. 1, SPAM free e-mail delivery may be achieved by simple operations achieved at a single subscriber's e-mail reception center by deleting SPAM from the e-mail stream addressed to that subscriber. As the e-mail addressed to this particular subscriber flows through the e-mail flow path from Internet Operations 15 to the e-mail subscriber server 17, SPAM is detected at the local SPAM Detector Station 16, which typically also stores the signature identification codes of currently active bulk mail messages. As indicated by the cable notation 19, the current messages passing through the Comparator Rejector Station 18 are compared with the stored SPAM signature codes and if a match is made, the SPAM is deleted, so that only SPAM-free e-mail is delivered into the subscriber's e-mail register 20.

Detailed Description Text (4):
In a similar manner shown in FIG. 2, the SPAM processing system may be located at the site of the Central Server 21, where SPAM free e-mail is made available at line 21 and sent to appropriate subscribers. This modified Comparator Spam-Remover 18' is programmed to remove the SPAM only from those subscribers requesting the service, except when it is in the interest of the Central Server 16 to reduce congestion and traffic by prohibiting the passage of all detected bulk mail which has not been licensed to use the services of that Central Server 16.

Detailed Description Text (5):
Thus, e-mail messages are processed in transit and the presence of SPAM is detected when identical messages, stripped of address and personalization data, are observed in a set of at least two, typically three, individual e-mail messages sent to different e-mail addresses. These messages are identified by a simple, easy to store, signature identification code, which is retained for further comparison, detection and processing of subsequent SPAM messages communicated in an e-mail flow path. The Central Server 16 may also via line 22 supply SPAM signatures to appropriate authorized users, typically having compatible e-mail comparator and SPAM rejection apparatus.

Detailed Description Text (10):
The in-line Signature Comparator 36 serves to compare the message signature of each e-mail message with the signatures in the SPAM Register 35 thereby to detect bulk mail currently being communicated. These messages may be processed either by attaching a SPAM ID flag 37, at least temporarily, to the message for later processing or to simply delete the message in the SPAM Deletion mechanism 38, thereby to produce the SPAMless e-mail flow stream 32.

Current US Original Classification (1):
709/206

CLAIMS:

7. The method of claim 5 further comprising the steps of processing an e-mail flow stream to determine identification signatures of messages flowing in said stream, and deleting those messages from the flow stream having signatures matching those identified from said different bulk mailings.

9. The method of claim 8 further comprising the step of deleting bulk mail identified by comparison of signatures from said sequence of e-mail messages with the stored signature codes of bulk mail messages.

23. The apparatus of claim 18 further comprising a signature code register for identified bulk mailings and comparator means for comparing signatures of signatures of e-mail messages in said flow stream with those of said register to identify bulk mailing messages.

☐ 6. Document ID: US 6321267 B1

L5: Entry 6 of 12 File: USPT Nov 20, 2001

DOCUMENT-IDENTIFIER: US 6321267 B1
TITLE: Method and apparatus for filtering junk email

Abstract Text (1):
An Active Filtering proxy filters electronic junk mail (also known as spam, bulk mail, or advertising) received at a Message Transfer Agent from remote Internet hosts using the Simple Mail Transfer Protocol (SMTP). The proxy actively probes remote hosts that attempt to send mail to the protected mail server in order to identify dialup PCs, open relays, and forged email. The system provides multiple layers of defense including: connect-time filtering based on IP address, identification of dialup PCs attempting to send mail, testing for permissive (open) relays, testing for validity of the sender's address, and message header filtering. A sender's message must successfully pass through all relevant layers, or it is rejected and logged. Subsequent filters feed IP addresses back to the IP filtering mechanism, so subsequent mail from the same host can be easily blocked.

Brief Summary Text (42):
The difficulty in filtering relayed junk mail is shown in part by this example. If the spammer 1060 forges the MAIL From address to match the relay host (e.g., "good@relay.dom") then as observed by net1.dom 1062, the message appears to be from a legitimate user at relay.dom. This example shows abuse of one open relay. The current generation of relaying tools will also permit the spammer to enter a list of open relay hosts, and the software will use different relays for different groups of addresses. Thus, different users at the same target network may receive spam relayed via different paths.

Brief Summary Text (52):
The solutions that are presently available to block junk mail fall into seven general categories. First, the use of centralized blacklisting databases, such as described above for the RBL, IMRSS, and DUL. Second, the use of local blacklisting databases, such as sendmail checking a local database and blocking email that matches entries in the database. Third, blocking mail from nonexistent domains, such

as for instance if sendmail receives "MAIL From: <sender@nonexistent.dom>", it will reject the mail because it cannot find the domain "nonexistent.dom" listed in the Domain Name System (DNS).

Brief Summary Text (61):
Minimal involvement is required by email administrators, when compared with the administrative cost of removing junk mail from mail servers, cleaning up after a virus or other malicious code attack, complaining about junk mail, and solving other problems. Administrator involvement generally consists of reviewing logs and adding IP address blocks and domain names to trusted databases where necessary.

Detailed Description Text (29):
Taken together, the processing performed by the Active Filtering proxy 1401 involves the following actions when a remote host 1400 establishes a TCP connection 1403 to the proxy. First, as shown at step 1406, the proxy server 1401 gets the IP address of the remote host and compares the IP address with a database of disallowed addresses. If the IP address of the remote host 1400 matches an entry in the database, the proxy server closes the TCP connection 1403 without transferring an email message. This is described in greater detail in FIG. 14.

Detailed Description Text (45):
At step 1405, the proxy determines if the remote host is categorized as trusted. Trusted networks are usually defined manually by using a suitable editor to enter IP addresses of trusted networks into the trusted database 1093 (FIG. 7). The proxy looks up the host name or IP address in a database of trusted network names and IP address blocks. This database is preferably a single linear file. A host name, e.g., "host37.remote.dom" matches an entry "remote.dom" if the two strings match from the last byte forward, for the length of the shorter string. If the host is trusted, processing continues with display of the greeting message in step 1409.

Detailed Description Text (46):
If the host is not trusted, the system proceeds to step 1406, which is also shown in FIG. 13. Here, the proxy determines whether the remote network has been blacklisted. The proxy compares the IP address of the remote host 1400 with entries in the blacklist database. Preferably, the blacklist database is implemented as a linear file containing one filter per line. Each filter consists of an ASCII dotted-quad address followed by a forward slash "/" and the number of bits to be compared, for example, "192.168.200.201/24", with optional textual information such as the date the filter was created, the host name, and the reason. The proxy compares the remote host's IP address with a filter entry by converting the two IP addresses to 32-bit values, XORs the two values, and right shifts the result so that only the specified number of bits (e.g., 24) remain. If the result is zero, then the remote host 1400 matches that particular filter.

Detailed Description Text (47):
At step 1408, if the remote host is blacklisted, the proxy 1401 issues an error reply to the remote host (e.g., "550 SMTP administratively blocked"), closes the connection 1403, logs the rejected connection, and exits without any email being transferred. The system log 1099 (FIG. 7) may be configured to log on the local host or on a remote host, such as the local MTA 1402. If the remote host 1400 is trusted or the IP address acquired in 1404 does not match any entry in the blacklist 1406, then the Active Filter displays the SMTP greeting message, step 1409.

Detailed Description Text (57):
At step 1417, the proxy filter 1401 skips subsequent checking of the MAIL From argument if either the connecting hostname or the MAIL From domain matches a trusted database entry, using the same method as in step 1405. The trusted database identifies networks with which there are long-term trust relationships, so that any user from one of these domains can send mail without restriction. If the domain is trusted, step 1417, processing continues with step 1470. Also at step 1417, the filter skips subsequent checking of the MAIL From argument if the MAIL From address (user@domain) exactly matches an entry in the whitelist database.

Detailed Description Text (58):
In the preferred embodiment, the whitelist file is a text file that contains

addresses (one per line) that are periodically mined from sendmail log entries for outgoing ("to=") messages. These log entries are for mail sent by the local organization to destination addresses on other networks, so adding these destination addresses to the whitelist file will ensure that the proxy will permit incoming email from those persons that local users have sent mail to. However, the whitelist database may be implemented as a hashed database (e.g., dbm) files, or even could be disabled. If the address matches a whitelist entry, processing continues with step 1470. The difference between the trusted database 1093 and the whitelist database 1094 is that for trusted hosts, mail is permitted from any user on the remote host to any user on the local host. For whitelist entries, mail is permitted only from the named user on the remote host to any user on the local host.

Detailed Description Text (59):
If the incoming connection is not itself a relay test and the message does not match any of the trust criteria, then in step 1418, the proxy 1401 attempts to open the reverse test connection to the remote host 1400. This is typically performed by calling socket( ) to acquire a socket structure to manage the connection to port 25 of the remote IP address 1400, then calling connect( ) to request the networking software to establish a TCP connection using the socket.

Detailed Description Text (65):
At step 1421, the proxy 1401 attempts to determine if the IP address or domain name matches a non-dialup entry in the dialup database. The Dialup configuration database 1097 (FIG. 7) lists blocks of non-dialup addresses that otherwise meet the criteria for dialup (i.e., will not accept a reverse connection and have a sequential naming scheme) but that are known to not be dialups. For example, an ISP may have sequentially-named mailhosts with some mailhosts dedicated for outgoing mail and some mailhosts dedicated for incoming mail.

Detailed Description Text (67):
The addresses in step 1421 are preferably expressed as a dotted-quad IP address, a slash "/", and a number of bits to be matched. For example, the filter 192.168.200.201/24 matches all addresses between 192.168.200.0 and 192.168.200.255. An address matches a particular filter if the filter address 1097 (FIG. 7) and the remote host 1400 IP address match for the specified number of bits. For example, the IP address 192.168.200.29 matches the filter 192.168.200.201/24 because the two addresses are identical for the first 24 bits, i.e., 192.168.200.

Detailed Description Text (68):
The preferred embodiment uses a flat ASCII file structure for the dialup database. If the requirement for non-dialup entries grow significantly, other representations (hashed lists, dbm files, or CAM) may be desirable for performance reasons. If the IP address matches any entry, then the proxy 1401 bypasses any further dialup testing, and proceeds to step 1901. Relay testing is not conducted since the filter has already determined that the reverse connection cannot be established to the remote host at step 1419. If it does not match any entry in the non-dialup list, then it proceeds with dialup testing in step 1422.

Detailed Description Text (69):
At steps 1422-1424, the proxy 1401 compares the name of the connecting host with its immediate neighbors, using a heuristic approach to correlate a sequence of names as dialups or non-dialups. In the preferred embodiment, a threshold total of ten match points are required to classify a remote host as a dialup. This approach takes into account the remote host name, character sequences in the name, and sequential nature of host names near the IP address of the remote host 1400.

Detailed Description Text (73):
At step 1423, the proxy 1401 compares the node name of the remote host 1400 with its neighbors and assigns additional points if the names appear to follow a sequential naming scheme. Further to the preferred embodiment, the proxy compares names of neighbor hosts by performing the following actions for all IP addresses that are within the range nnn-10 to nm+10, where nnn is the node address (last byte of IP address) of the remote host 1400. Details of step 1423 are provided in FIG. 17. For example, the following Table 3 for a remote host 1Cust117.tnt5.salt-lake-city.ut.da.uu.net at IP address 63.11.217.117, shows the IP

addresses and node names for its 20 nearest neighbors.

Detailed Description Text (75):
At step 1424, the proxy 1401 compares the total current number of match points from steps 1422 and 1423 with the threshold number of points (10, in the preferred embodiment) required to characterize the remote host as a dialup. If the number of match points exceeds the threshold, then it exits, step 1425. Otherwise, message transfer continues with step 1901.

Detailed Description Text (77):
FIG. 17 shows further detail of the processing flow for step 1423 in accordance with the preferred embodiment. Step 1500 calculates a 32-bit IP address for the remote host, which is used in step 1504 to calculate the IP address of one of its 20 neighbors. Steps 1501, 1502, and 1503 perform the remaining steps shown in the figure for x=-10 to x32 +10, inclusive, while skipping the remote host at x=0. When the loop is finished, the proxy exits to step 1424 of FIG. 16, which classifies the remote host as a dialup or non-dialup based on the accumulated number of match points.

Detailed Description Text (78):
Step 1505 limits the name comparison to the 8-bit (Class C) address block that contains the remote host to avoid comparing a remote host name in one ISP with neighbors in a block operated by a different ISP. It XORs the 32-bit IP address for the neighbor x and the IP address for the remote host and shifts the result right 8 bits. If the result is non-zero, then the neighbor x is in a different address block than the remote host, and is skipped. Thus, the range is absolutely bounded by a minimum node address of 0 and a maximum node address of 255, so that the comparison for remote host 192.168.200.2 would only consider node addresses from 0-1 and 3-12, in order to avoid comparing names in other 8-bit blocks of addresses.

Detailed Description Text (79):
At a minimum, ten addresses will always be considered. Preferably, 10 names are matched out of 20 (n-10 to n+10) so that if the remote host is at the beginning of a block, e.g., 192.135.140.0, then there will still be ten opportunities to match from 1 to 10, and if the remote host is at the end of a block (e.g., 192.135.140.255, there will still be ten opportunities to match in the range 245 to 254.

Detailed Description Text (80):
Steps 1506 and 1507 call gethostbyaddr() to get the host structure for the neighbor x, which contains the host name. Errors do not terminate the comparison, since there may be gaps in the DNS information near the remote host. Steps 1509 and 1510 compare the respective lengths of the remote host name and its neighbor x. If either is more than one character longer than the other, then skip the neighbor x because the two names do not appear to be part of a sequence.

Detailed Description Text (81):
Step 1511 scans forward and backwards to identify the sequence of non-matching characters in the names of the remote host and its neighbor x. This sequence may contain substrings of matching characters, but as shown in step 1512, if either string is greater than three characters in length, then skip the neighbor x because the two names do not appear to be part of a sequence.

Detailed Description Text (83):
In step 1517-1519, the proxy calculates the absolute distance between the two name sequences. If the distance is less than or equal to the absolute value of x, then the names appear to be part of a sequence and the match counter is incremented. For example, Table 4 shows the distance as correlated to the offset x for the four nearest neighbors of the remote host 63.11.217.117, based on the information in Table 3. As shown in this example, the distance calculated for each of the four nearest neighbors is identically equal to the difference in IP address values, thus the names are part of a sequence.

Detailed Description Text (85):
The preferred embodiment supplements these Active Dialup methods with the blacklist filter 1406 (FIG. 14). Bulk mailers who use the SMTP direct mechanism will typically

retry from different (dynamically assigned) IP addresses, but frequently from addresses in the same Class C (8-bit) address range. By adding the IP address to the blacklist database, nominally with the number of bits to be matched set to 24, the mechanism takes advantage of the relative speed of a blacklist database lookup as compared with subsequent iterations of the above active dialup mechanism. The filter can be left perpetually in the blacklist database, or preferably removed from the blacklist database if it is not used in some number of days or weeks. Further, filters can be manually added to handle any blocks of dialup addresses that are not identified by the Active Dialup method.

Detailed Description Text (91):
Possible threshold values are 1, 2, and 3, since an edit distance of 0 would indicate identity and thus not be useful and an edit distance greater than 3 is too broad and would result in miscategorization of non-dialup hosts as dialups. An edit distance of 1 would indicate a high degree of correlation (names would match only if they differed by one character), but would fail to match names such as "dial39" and "dial40". A threshold of 2 is probably optimal, even though it would improperly categorize rollover situations such as "dial99" to "dial100". A threshold of 3 would address the aforementioned rollover problem, but would miscategorize, for example, a remote host "mail" surrounded by a sufficient number of hosts with names such as "main", "mail2", and "menu". In step 1531, the proxy scans each character of the neighbor name and compares it with the corresponding character in the remote host name. If the two characters are identical (step 1532), then the proxy advances the character pointer in the two names. In steps 1533, 1534, and 1535, the proxy determines if a character must be replaced, inserted, or deleted to make the neighbor name consistent with the remote host name. If so, it increments the edit distance 1536 for the neighbor name and continues. When the comparison is complete, the proxy then checks if the accumulated edit distance is less then or equal to the threshold read from the Dialup DB in step 1530. If so, it increments the match count 1538. The proxy then continues with the next name, as determined by step 1502.

Detailed Description Text (93):
Still another alternative embodiment involves categorizing the remote host based on the ability to establish reverse test connections to its neighbors as well as the remote host itself, step 1418 (FIG. 15). If a sufficient number of neighboring addresses also do not permit reverse test connections, then it is reasonable to conclude that the remote host is a dialup. This method might be used by itself to replace the method shown in FIG. 17, or may be combined with the method in FIG. 17. With respect to FIG. 23, the steps 1500-1520 are as described in FIG. 17 and provide a means of stepping through each of the 20 nearest IP addresses for the remote host. In step 1550, the proxy attempts to connect to the neighbor x, using the same means described in step 1418. It checks the status for the connection in step 1551. If the connection is not successful 1552, as would be typical for a block of dialup addresses, the proxy increments the match count. However, if the proxy is able to establish a reverse connection to the neighbor x, it subtracts 2 from the match count 1553 and closes the test connection 1554. This weighting permits as many as three neighbors to accept reverse connections and still categorize the remote host as a dialup. The proxy then continues with the next name, as determined by step 1502. The disadvantage of this approach is that it may be time-consuming to attempt a large number of reverse SMTP connections. However, it is less time consuming to perform this test in SMTP filtering software than it is to deal with the spam or junk mail after it is received on the organization's mail server.

Detailed Description Text (108):
At step 1462, the proxy 1401 attempts to find if the IP address of the remote host 1400 matches a non-relay entry in the Relay database 1096 (FIG. 7). This database lists blocks of addresses that the local organization must exchange email with, but which would fail the relay test. There might typically be between about 5-50 entries in this database, with each entry covering a block of addresses. These entries can be pre-defined by a site survey performed by each organization, preferably before installing the Active Filtering proxy server. For simplicity, the preferred embodiment of the Relay Database 1096 (as with other IP addresses listed in steps 1406 and 1413) expresses these addresses as a dotted-quad IP address, a forward slash "/", and a number of bits to be matched. Other embodiments may use other representations (hashed lists, dbm filed, or CAM) for performance reasons.

Detailed Description Text (109):
If the IP address matches any non-relay entry, then the proxy 1401 bypasses any further relay and user testing, and proceeds with message transfer in step 1470. If the IP address of the remote host does not match any entry in the non-relay list, then it continues with the second part of Active Relay testing at step 1463. Before performing the relay test, the proxy compares the MAIL From domain with the connecting host name. A match at step 1463 occurs when the connecting host name and the MAIL From domain are identical, beginning at the end of the two strings, and comparing backwards over the last two nodes (i.e., periods) of the two domains. For example, if host "smtp.gamma.dom" connects with MAIL From "alpha@gamma.dom", the two domains match over the scope "gamma.dom". However, if host "smtp.gamma.dom" connects with MAIL From "alpha@beta.dom", the two domains do not match.

Detailed Description Text (110):
If the two domains match in step 1463, the proxy checks the Relay database 1096 at step 1464 to determine if an administrator has configured the proxy to perform loose relay testing. With loose testing, the proxy permits mail from open relays if the MAIL From address matches the connecting host name. In this case, the relay test message 1465 is not necessary, so the proxy continues with transfer of the message in step 1470.

Detailed Description Text (111):
If either of the domains do not match in step 1463 or the Relay database is configured for strict relay testing, the proxy issues a RCPT message 1465 identifying a configurable string, defined when the proxy is installed. The default includes the name "relayto" and the local domain name (for legal reasons), for example, RCPT To: <relayto@local.dom>. The configurable recipient address may be any syntactically correct address. Even though a message is not sent the RCPT address is preferably not a real user address in order to avoid address mining by spam site administrators.

Detailed Description Text (122):
The preferred embodiment of the Active Relay method performs testing in the order shown in FIG. 18. In particular, this ordering avoids the relay test sequence if the remote host is configured as a non-relay (step 1462) or the domains match and loose relay testing is configured (steps 1463 and 1464). However, alternative embodiments may provide for other orders of testing.

Detailed Description Text (133):
Preferably, when appending an IP address to the blacklist database 1095, the proxy adds a 4-byte IP address, e.g., 192.168.200.45, along with some number of bits to be matched. Typically the number of bits is 24, so that any subsequent connection by any host in the range 192.168.200.0 through 192.168.200.255 will be rejected by the blacklist mechanism. This takes into account that Class C addresses are normally assigned to organizations in multiples of 256, so subsequent connections in the 192.168.200.x range are normally owned by the same irresponsible organization, so it makes sense to block all of them. However, if the ownership is subsequently determined to be something more or less than a single class C, then the blacklist file can be manually edited to block one or more hosts.

Detailed Description Text (137):
In step 1470 the proxy connects to the MTA using the same method described for the reverse test connection 1418. In summation, the proxy connects to the MTA for messages with any of the following characteristics: connection from a trusted domain (step 1417); trusted MAIL From domain (step 1417); whitelisted MAIL From address (step 1417); or email from a user with an account at a non-dialup, non-relay host (step 1467). In addition, subject to the validity of the user address as determined in step 1913, the proxy will permit mail from hosts where the reverse connection fails, but host is configured as non-dialup (step 1421); reverse connection fails, but host is not detected as a dialup (step 1424); reverse connection succeeds, but relay test is inconclusive (steps 1454, 1456, 1458); reverse connection succeeds, but host is configured as non-relay (step 1462) or reverse connection succeeds, MAIL From address matches connecting host, and the proxy is configured for loose relay testing (steps 1463, 1464).

Current US Original Classification (1):
709/229

Current US Cross Reference Classification (2):
709/218

Current US Cross Reference Classification (3):
709/227

Current US Cross Reference Classification (4):
709/238

☐ 7. Document ID: US 6161130 A

L5: Entry 7 of 12 File: USPT Dec 12, 2000

DOCUMENT-IDENTIFIER: US 6161130 A
**** See image for Certificate of Correction ****
TITLE: Technique which utilizes a probabilistic classifier to detect "junk" e-mail by automatically updating a training and re-training the classifier based on the updated training set

Abstract Text (1):
A technique, specifically a method and apparatus that implements the method, which through a probabilistic classifier (370) and, for a given recipient, detects electronic mail (e-mail) messages, in an incoming message stream, which that recipient is likely to consider "junk". Specifically, the invention discriminates message content for that recipient, through a probabilistic classifier (e.g., a support vector machine) trained on prior content classifications. Through a resulting quantitative probability measure, i.e., an output confidence level, produced by the classifier for each message and subsequently compared against a predefined threshold, that message is classified as either, e.g., spam or legitimate mail, and, e.g., then stored in a corresponding folder (223, 227) for subsequent retrieval by and display to the recipient. Based on the probability measure, the message can alternatively be classified into one of a number of different folders, depicted in a pre-defined visually distinctive manner or simply discarded in its entirety.

Brief Summary Text (22):
In accordance with our specific inventive teachings, each incoming e-mail message, in such a stream, is first analyzed to determine which feature(s) in a set of N predefined features, i.e., distinctions, (where N is an integer), that are particularly characteristic of spam, the message contains. These features (i.e., the "feature set") include both simple-word-based features and handcrafted features. A feature vector, with one element for each feature in the set, is produced for each such message. The contents of the vector are applied as input to a probabilistic classifier, such a modified Support Vector Machine (SVM) classifier, which, based on the features that are present or absent from the message, generates a continuous probabilistic measure as to whether that message is spam or not. This measure is then compared against a preset threshold value. If, for any message, its associated probabilistic measure equals or exceeds the threshold, then this message is classified as spam and, e.g., stored in a spam folder. Conversely, if the probabilistic measure for this message is less than the threshold, then the message is classified as legitimate and hence, e.g., stored in a legitimate mail folder. The contents of the legitimate mail folder are then displayed by a client e-mail program for user selection and review. The contents of the spam folder will only be

displayed by the client e-mail program upon a specific user request. The messages in the spam folder can be sorted by increasing probability that the messages are spam, so that the user need only check that the top few messages are indeed spam before deleting all the messages in the folder.

Detailed Description Text (15):
A feature vector, with one element for each feature in the set, is produced for each incoming e-mail message. That element simply stores a binary value specifying whether the corresponding feature is present or not in that message. The vector can be stored in a sparse format (e.g., a list of the positive features only). The contents of the vector are applied as input to a probabilistic classifier, preferably a modified support vector machine (SVM) classifier, which, based on the features that are present or absent from the message, generates a probabilistic measure as to whether that message is spam or not. This measure is then compared against a preset threshold value. If, for any message, its associated probabilistic measure equals or exceeds the threshold, then this message is classified as spam and, e.g., stored in a spam folder. Alternatively, if the probabilistic measure for this message is less than the threshold, then the message is classified as legitimate and hence, e.g., stored in a legitimate mail folder. The classification of each message is also stored as a separate field in the vector for that message. The contents of the legitimate mail folder are then displayed by the client e-mail program for user selection and review. The contents of the spam folder will only be displayed by the client e-mail program upon a specific user request.

Detailed Description Text (21):
As shown, the software modules utilized in this embodiment of our invention include: mail classifier 210 that itself contains: handcrafted feature detector 320, text analyzer 330, indexer 340, matrix/vector generator 350, feature reducer 360, classifier 370 and threshold comparator 380; and mail store 220.

Detailed Description Text (29):
In essence, during the classification phase, each incoming e-mail message is quantitatively classified, through classifier 370, to yield an output confidence level which specifies a probability (likelihood) that this particular message is spam. As noted above, this likelihood can be used to drive several alternative user-interface conventions employed to allow review and manipulation of spam. For a binary folder threshold-based approach to displaying and manipulating the spam probability assignment, the message is designated as either spam or legitimate e-mail, based on the magnitude of the assigned probability of spam, and then illustratively stored in either spam folder 227 or legitimate mail folder 223, respectively, for later retrieval. To do so, each incoming message is first analyzed but only to detect the presence of every individual handcrafted and word-oriented feature in the N-element feature set, thereby resulting in an N-element feature vector for that message. This vector is applied as input to the classifier. In response, the classifier produces an output confidence level, i.e., a classification probability (likelihood), for that particular message. The value of this probability is compared against a fixed threshold value. Depending upon whether the probability equals or exceeds, or is less than the threshold, the message is classified as either spam or legitimate mail and is then stored in the corresponding mail folder.

Detailed Description Text (31):
Given this vector, classifier 370 then generates an associated quantitative output confidence level, specifically a classification probability, that this particular message is spam. This classification probability is applied, as symbolized by line 375, to one input of threshold comparator 380. This comparator compares this probability for the input message against a predetermined threshold probability, illustratively 0.999, associated with spam. If the classification probability is greater than or equal to the threshold, then the input message is designated as spam; if the classification probability is less than the threshold, then this input message is designated as legitimate mail. Accordingly, the results of the comparison are applied, as symbolized by line 385, to mail store 220 to select a specific folder into which this input message is then to be stored. This same message is also applied, as symbolized by line 205 and in the form received, to an input of mail store 220 (this operation can be implemented by simply accessing this message from a common mail input buffer or mail queue). Based on the results of the comparison, if

this message is designated as legitimate mail or spam, it is then stored, by mail store 220, into either folder 223 or 227, respectively. The legitimate mail and spam can be rank ordered within their respective folders 223 or 227 in terms of their corresponding output confidence levels. In this regard, e.g., the legitimate mail could be ordered in legitimate mail folder 223 in terms of their ascending corresponding confidence levels (with the messages having the lowest confidence levels being viewed by the classifier as "most" legitimate and hence being displayed to a recipient at a top of a displayed list, followed by messages so viewed as having increasingly less "legitimacy"). Furthermore, not only can these messages be rank ordered but additionally or alternatively the messages themselves (or portions thereof) or a certain visual identifier(s) for each such message can be color coded. Specific colors or a range of colors could be used to designate increasing levels of legitimacy. In this case, a continuous range (gamut) of colors could be appropriately scaled to match a range that occurs in the output confidence level for all the legitimate messages. Alternatively, certain predefined portions of the range in the output confidence level could be assigned to denote certain classes of "legitimacy". For example, a red identifier (or other color that is highly conspicuous) could be assigned to a group of mail messages that is viewed by the classifier as being the most legitimate. Such color-coding or rank ordering could also be incorporated as a user-controlled option within the client e-mail program such that the user could customize the graphical depiction and arrangement of his(her) mail, as desired. Furthermore, such color coding can also be used to denote certain categories of spam, e.g., "certain spam", "questionable spam" and so forth. Alternatively, other actions could occur, such as outright deletion of a message, based on its classification probability, e.g., if that probability exceeds a sufficiently high value.

Detailed Description Text (35):
Classifier 370 can be implemented using a number of different techniques. In that regard, classifier 370 can be implemented through, e.g., a support vector machine (SVM) as will be discussed in detail below, a Naive Bayesian classifier, a limited dependence Bayesian classifier, a Bayesian network classifier, a decision tree, content matching, neural networks, or any other statistical or probabilistic-based classification technique. In addition, classifier 370 can be implemented with multiple classifiers. Specifically, with multiple classifiers, each such classifier can utilize a different one of these classification techniques with an appropriate mechanism also being used to combine, arbitrate and/or select among the results of each of the classifiers to generate an appropriate output confidence level. Furthermore, all these classifiers can be the same but with, through "boosting", their outputs weighted differently to form the output confidence level. Moreover, with multiple classifiers, one of these classifiers could also feed its probabilistic classification output, as a single input, to another of these classifiers.

Detailed Description Text (87):
By using the target values ti rather than y.sub.i, the resulting sigmoid function is no more precise than the data used to determine it. That is, the sigmoid function is not "overfit" to past data, particularly in categories with little training data. In this way, a sigmoid function is determined that matches unknown data given a prior that all probabilities are equally likely. Although the use of the target function t.sub.i of equation (18) is presented below in the context of creating classifier 370, it is applicable to other classifiers as well. Naturally, other target functions, which depend on the number of training examples in the positive and negative classes, could be used instead of the target function defined in equations (17).

Detailed Description Text (112):
Execution next proceeds to decision block 650, which effectively implements using, e.g., a sigmoid function as described above, threshold comparator 380. In particular, block 650 determines whether the classification probability of message j, i.e., p.sub.j, is greater than or equal to the predefined threshold probability, p.sub.t, (illustratively 0.999) for spam. If, for message j, its classification probability is less than the threshold probability, then this message is deemed to be legitimate. In this case, decision block 650 directs execution, via NO path 653, to block 660. This latter block, when executed, sets the classification for incoming

message j to legitimate and stores this classification within an appropriate classification field in the feature vector for this message. Thereafter, execution proceeds to block 670 which stores message j within legitimate mail folder 223 for subsequent retrieval by and display to its recipient. Once this message is so stored, execution exits, via path 675 from process 600. Alternatively, if, for message j, its classification probability exceeds or equals the threshold probability, then this message is deemed to be spam. In this case, decision block 650 directs execution, via YES path 657, to block 680. This latter block, when executed, sets the classification for incoming message j to spam and stores this classification within an appropriate classification field in the feature vector for this message. Thereafter, execution proceeds to block 690 which stores message j within spam folder 227 for possible subsequent retrieval by and/or display to its recipient. Once message j is stored in this folder, execution then exits, via path 695 from process 600.

Current US Original Classification (1):
709/206

Current US Cross Reference Classification (3):
709/205

Current US Cross Reference Classification (4):
709/207

Current US Cross Reference Classification (5):
709/240

Other Reference Publication (9):
Hinrich Schutze et al, "A Comparison of Classifiers and Document Representations for the Routing Problem", Proceedings of the 18.sup.th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, Jul. 9-13, 1995, pp. 229-237.

Other Reference Publication (10):
Yiming Yang et al, "A Comparative Study on Feature Selection in Text Categorization", School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, and Verity, Inc., Sunnyvale, CA.

Other Reference Publication (12):
David D. Lewis et al, "A Comparison of Two Learning Algorithms for Text Categorization", Third Annual Symposium on Document Analysis and Information Retrieval, Las Vegas, Nevada, Apr. 11-13, 1994, pp. 81-93.

CLAIMS:

4. The method in claim 2 further comprising the steps of:

comparing the output confidence level for the incoming message to a predefined probabilistic threshold value so as to yield a comparison result; and

distinguishing said incoming message, in a predefined manner associated with the first class, from messages associated with the second class if the comparison result indicates that the output confidence level equals or exceeds the threshold level.

17. The method in claim 8 wherein the probabilistic classifier comprises a Naive Bayesian classifier, a limited dependence Bayesian classifier, a Bayesian network classifier, a decision tree, a support vector machine, or is implemented through use of content matching.

18. The method in claim 17 wherein:

the feature data applying step comprises the step of yielding the output confidence level for said incoming message through a support vector machine; and

the comparing step comprises the step of thresholding the output confidence level

through a predefined sigmoid function to produce the comparison result for the incoming message.

37. The apparatus in claim 35 wherein the processor, in response to the stored instructions:

compares the output confidence level for the incoming message to a predefined probabilistic threshold value so as to yield a comparison result; and

distinguishes said incoming message, in a predefined manner associated with the first class, from messages associated with the second class if the comparison result indicates that the output confidence level equals or exceeds the threshold level.

50. The apparatus in claim 41 wherein the probabilistic classifier comprises a Naive Bayesian classifier, a limited dependence Bayesian classifier, a Bayesian network classifier, a decision tree, a support vector machine, or is implemented through use of content matching.

51. The apparatus in claim 50 wherein the processor, in response to the stored instructions:

yields the output confidence level for said incoming message through a support vector machine; and

thresholds the output confidence level through a predefined sigmoid function to produce the comparison result for the incoming message.

☐ 8. Document ID: US 6052709 A

L5: Entry 8 of 12 File: USPT Apr 18, 2000

DOCUMENT-IDENTIFIER: US 6052709 A
TITLE: Apparatus and method for controlling delivery of unsolicited electronic mail

Abstract Text (1):
In a system and method and system for controlling delivery of unsolicited electronic mail messages, one or more spam probe e-mail addresses are created and planted at various sites on the communications network in order to insure their inclusion on large-scale electronic junk mail ("spam") mailing lists. The mailboxes corresponding to the spam probe e-mail addresses are monitored for incoming mail by a spam control center. Upon receipt of incoming mail addressed to the spam probe addresses, the spam control center automatically analyzes the received spam e-mail to identify the source of the message, extracts the spam source data from the message, and generates an alert signal containing the spam source data. This alert signal is broadcast to all network servers and/or all user terminals within the communications network. A filtering system implemented at the servers and/or user terminals receives the alert signal, updates stored filtering data using the spam source data retrieved from the alert signal, and controls delivery of subsequently-received e-mail messages received from the identified spam source. The filtering system controls delivery of the unsolicited e-mail messages by discarding the messages without displaying them to the user, displaying the messages to the user with a "JUNK" or similar marker, or otherwise processing the spam mail as desired by the network provider and/or the network users. The filtering system may also filter e-mail messages sent by the user terminals.

Detailed Description Text (12):
Filtering instructions may be transmitted from the control center to the servers or user terminals or both to update and control the filtering operation performed by the servers or user terminals or both. For example, a filtering instruction may instruct the server or user terminal to compare the domain of the MESSAGE ID with the earliest source in the series of servers listed in the received field. If the domains of the two fields do not match, the server or user terminal would mark the e-mail message with a code, such as JUNK, indicating its status as spam. Other appropriate filtering instructions as would be apparent to one of skill in the art may also be provided in order to insure effective spam filtering.

Detailed Description Text (17):
Filtering of incoming e-mails may preferably be performed as follows. If the data in any of the "FROM" field of the incoming e-mail message match data stored in the corresponding data category of the exclusion list manager 202, the e-mail is marked by the filter 204 with a first display code indicating the "JUNK" status of the message. Optionally, the filter 204 may use multiple display codes indicating multiple status levels of "JUNK." If no match is detected between the fields of the received e-mail message and the stored exclusion list, the e-mail is marked with a second display code indicating the "OK" status of the message. Notably, the filtering system may be programmed to search not only for precise text in matching data fields, but also for similar text using known text searching techniques.

Detailed Description Text (20):
In another preferred embodiment of the filtering system according to the present invention, e-mail messages marked with the first display code ("JUNK" mail) are further processed by the filter using user preference data entered by the user. The user may, for example, desire to receive unsolicited e-mail messages relating only to one or more specific subjects. Accordingly, the user may enter into his or her terminal a list of subjects which is stored as preference data by the filtering application. The filter application compares the subject data of the received e-mail message with subject preference data entered by the user. Notably, the subject data from the received message may include "SUBJECT" header information, the full text of the e-mail message, or both. A text search may be performed to determine whether the subject data from the received e-mail contains any of the subject words or phrases entered by the user as preference data. If a match is detected, the e-mail message is marked with a third display code and displayed to the user in a third distinctive mode using known display techniques. These e-mail messages may, for example, be automatically placed in a special folder created by the user or the filtering application or displayed in a distinctive color.

Detailed Description Text (21):
In yet another preferred embodiment of a filtering application for use in the present invention, the filtering application stores predetermined classification data, which are used to sort incoming unsolicited e-mail messages into predetermined categories. The filtering application may, for example, search the "SUBJECT" header data of the received e-mail message, other headers, the text of the message, or all three, to determine whether the subject data from the received message contain any words or phrases matching the subject information describing each predetermined category. In this embodiment, each predetermined subject category of messages is associated with a specific display code. Accordingly, received messages in each predetermined category would be displayed to the user in different display modes to visually distinguish the categories on the user's display screen. The user may select to receive unsolicited e-mail messages in one or more of the predetermined categories, or none of the categories.

Detailed Description Text (27):
The filtering process performed for each user terminal, e.g., 130, by the e-mail filter 504 is the same as that performed by filter 204 in FIG. 2. The filter 504 compares the data stored in the "FROM" header field (and optionally the "TO," "CC," "BCC," and "SUBJECT" fields and associated sub-headers fields) of the incoming e-mail messages with corresponding categories of data stored in the exclusion list processor 502. If data in any of these fields of the incoming email matches data stored in a corresponding field of the inclusion list processor 502, the incoming e-mail is marked "JUNK" and marked with a first display code. If no match is

detected, the e-mail filter labels the e-mail message as "JUNK" by marking the message with a second display code. Further processing to display the JUNK messages by subjects or subject categories as described above with reference to FIG. 5 may also be performed by the server's filtering application.

Current US Original Classification (1):
709/202

Current US Cross Reference Classification (1):
709/204

☐ 9. Document ID: US 6023723 A

L5: Entry 9 of 12 File: USPT Feb 8, 2000

DOCUMENT-IDENTIFIER: US 6023723 A
TITLE: Method and system for filtering unwanted junk e-mail utilizing a plurality of filtering mechanisms

Abstract Text (1):
A system and method of filtering junk e-mails. The user is provided with or compiles a list of e-mail addresses or character strings which the user would not wish to receive to produce a first filter. A second filter is provided including names and character strings which the user wishes to receive. Any e-mail addresses or strings contained in the first filter will be automatically eliminated from the users system. Any e-mail addresses or strings contained in the second filter would be automatically sent to the users "In Box". Any e-mail not provided in either of the filter lists will be sent to a "Waiting Room" for user review. If this user review results in the user rejecting any e-mail, the address as well as specific character strings included in this e-mail would be transmitted to a central location to be included in a master list. This master list is periodically sent to each of the users allowing the first filter to be updated.

Detailed Description Text (3):
The software would allow the individual user to construct an automatic discard filter 12. The automatic discard filter is a collective term consisting of a user modified discard filter, a user personal address filter as well as a user personal string filter. During operation of the system, the automatic discard filter 12 would include a current filter list comprising a list of active e-mail addresses against which incoming e-mails are compared. This current filter list is retained in a memory section of the users computer. Any comparison between any incoming e-mail and the current filter list could be accomplished within the user's computer system. The current filter list is maintained at the remote central location 46 as well as being periodically updated in each of the users PC systems 48. The remote location 46 would include a delta filter server 22 and download server 24 for a particular user as well as delta server filter 26 from all other users. The current filter list can be modified by the user to personally remove any addresses therefrom through various deletion techniques, thereby providing the user with a user modified discard filter. The user personal address filter would include additional addresses the user has added to the current filter list as well as any character strings that the user has added via a text entry containing an "@". For the purpose of the present invention, a text entry is a character string entered into the system by keyboard typing. Typing is initiated by double clicking or highlighting and typing, thereby clearing an old string and creating a new string. When the mouse is clicked on some other location or "enter" is hit, the string will be entered into the appropriate memory structure for the new field.

Detailed Description Text (6):
Any e-mail received by the user is checked against the automatic discard filter 12 to determine whether any character string on the "No Admittance List" 52 will bar entry of any e-mail with matching text in its address, subject line or message body. If this occurs, that e-mail will be eliminated from the users system as indicated by the Trash Bin 16.

Detailed Description Text (7):
Conversely, any address contained in the Guest List which matches an address of an incoming e-mail would be automatically forwarded to the In Box folder 18 for review by the user. Similar to the situation with respect to the No Admittance List 52, a text string entered in the Guest List 54 would forward all messages containing that character string to the "In Box" folder 18. This feature would allow users to receive on-demand direct marketing information from parties, promoting products for which the user has expressed interest based upon the text string entered in the Guest List 54.

Detailed Description Text (10):
Periodically, the database server 24 in communication with the address filter server 22 would download updated filter addresses to the various users in the system by constructing an address packet consisting of every address on the current filter list since the date and time of each of the users last update. The address packet is a data structure consisting of N strings of e-mail addresses and a variable containing the time of construction of the packet. The packet is compressed for downloading and uploading multiple e-mail addresses. Based upon the particular implementation of the software of the present invention, the updated version of the current filter list is substituted for the No Admittance List currently provided in the users system. Alternatively, since the No Admittance List might include addresses and character strings personally added by the user but not included in the current filter list, the updated filter list would be compared with the automatic discard filter and any additional entries not included in the automatic discard filter would be added thereto.

Detailed Description Text (11):
FIG. 2 illustrates the In Box folder 18 and the Waiting Room 20 in more detail as well as giving examples of the type of messages included therein. The list of names included on the automatic discard filter 12 are provided in the No Admittance List 20. Any incoming e-mail whose new address matches one of the addresses on this list is immediately discarded to Trash 16. Addresses may be added to this list via an update button 61, the Add to No Admittance button 58, text entry, or by dragging a selected e-mail to this window with the mouse. The update button 62 automatically downloads the latest automatic discard filter from the download server 24. The updated filter list is displayed in the No Admittance Window. Simultaneously, user added e-mail addresses are sent to the Delta Filter Server 22 for consideration in future updates to the users in the system.

Detailed Description Text (15):
The Guest List window 54 would include a list of names on the Guest List filter. Any incoming e-mail whose new address matches one of the addresses on this list is immediately forwarded to the In Box folder 18. Addresses may be added to this list via the add to Guest List button 60, text entry, or by dragging a selected e-mail to this window with a mouse.

Detailed Description Text (16):
The In Box 18 includes only those e-mails that have successfully passed through both the automatic discard filter and the Guest List Filter. Additionally, any e-mail from any folder, may be selected and dragged into the In Box 18 by the user using the mouse. Similar to the Waiting Room 20 the In Box 18 includes the e-mail addresses, the date and time of receipt as well as the subject matter of the e-mail. Furthermore, the particular configuration of the In Box 18 as illustrated in FIG. 2 can be changed depending upon the users requirements. By clicking on an open slot in the No Admittance List 52 or the Guest List 54 or by double clicking on a existing text, the user may enter a character string to be checked in the filtering system. Any such character string on the No Admittance List will bar the entry of any e-mail with matching text in its address, subject line or message body. For example, as

shown in the No Admittance List 52, any received e-mail with the words "free money" in its subject or message would be discarded. A text string similarly entered in the Guest List would forward all messages containing that character string to the In Box. Text entry can also be used to type in new e-mail addresses or edit existing ones on either of the filter lists.

Detailed Description Text (22):
Both the control screen for the member database 62 and the address database 70 contain a search field 64, a search panel 66 and a filter panel 68. The search field 64 would contain information matching an entry in either the address database or the member database. Buttons 72 and 74 would allow either of these databases to appear on the control screen. Both of these databases would include search results run in either the address database or the member database in Section 66. The current filter Section 68 would allow entries to be updated or saved at various times. It would also include a box 76 indicating the number of days an address can remain on the current filter list without a new instance of that address being uploaded by the filter users. It would also include a box 78 listing the minimum number of reportings required for an address to be placed on the current filter list. Certainly both of these central screens can be set up in multitude of ways depending upon the specific information to be provided.

Detailed Description Text (23):
Returning to FIG. 1, the process of comparing received e-mails to both the Automatic Discard filter 12 and the Guest List filter 14 will now be explained. Incoming e-mails 28, 30 and 32 are compared to information contained in the user modified discard filter, the user personal address filter and the users personal string filter utilizing the address line, the subject line as well as the message body. Since the information included in e-mails 28, 30 and 32 are not contained in the automatic discard filter, all three of these e-mails are directly transmitted to the Guest List filter 14. The e-mail addresses, subject line and message body of these three e-mails result in a match for all three of these e-mails. Consequently, these e-mails are sent to the In Box folder 18.

Current US Original Classification (1):
709/206

Current US Cross Reference Classification (8):
709/203

Current US Cross Reference Classification (9):
709/207

Current US Cross Reference Classification (10):
709/231

Current US Cross Reference Classification (11):
709/238

Current US Cross Reference Classification (12):
709/240

Current US Cross Reference Classification (13):
709/242

Current US Cross Reference Classification (14):
709/249

CLAIMS:

1. A system for filtering electronic mail for a plurality of user computers received over a communication medium to each of the user computers comprising:

a first filter provided in each of the user computers provided with a first list of addresses from which the user does not wish to receive electronic mail;

a second filter provided in each of the user computers provided with a second list of addresses from which the user does wish to receive electronic mail;

comparison device provided in each of the user computers for comparing the addresses included in said first filter and said second filter with the address of electronic mail received by the user computers, wherein any electronic mail whose address is discovered by said comparison device to be included in said first list of addresses will be automatically deleted from that particular user computer, and further wherein any electronic mail whose address is discovered by said comparison device to be included in said second list of addresses would be retained in that particular user computer for review by the user;

a first folder for retaining electronic mail received in a user computer whose address is not included in either said first list of addresses or said second list of addresses for review by the user at a later time, and further including an update device for adding addresses to said first filter or said second filter based upon the user's review, thereby creating a first updated list of first addresses and a first updated list of second addresses;

a device for periodically sending said first updated list of first addresses from a plurality of user computers to a master database situated at a location remote from the user's computer thereby creating an updated master database of addresses received from the user computers; and

a device provided at said location remote from the user's computers for including addresses of said undated master database only if the same address is sent to said master database by a predetermined number of user computers.

5. A system for filtering electronic mail provided with a first list of addresses from which the user does not wish to receive electronic mail as well as a first list of objectionable character strings;

a second filter provided in each of the user computers provided with a second list of addresses from which the user does wish to receive electronic mail as well as a second list of acceptable character strings;

comparison device provided in each of the user computers for comparing the addresses and character strings included in said first filter and said second filter with the address of electronic mail as well as character strings provided in the electronic mail received by the user computers, wherein any electronic mail whose address is discovered by said comparison device to be included in said first list of addresses or electronic mail containing a character string included in the first list of objectionable characters string, will be automatically deleted from that particular user computer, and further wherein any electronic mail whose address is discovered by said comparison device to be included in said second list of addresses or said second list of acceptable character strings, would be retained in that particular user computer for review by the user;

a first folder for retaining electronic mail received in the computer whose address is not included in said first list of addresses, said first list of objectionable character strings, said second list of addresses or said second list of acceptable character strings, for review by the user at a later time, and further including an update device for adding addresses to said first filter or said second filter based upon the user's review, creating a first updated list of first addresses and a first updated list of second addresses;

a device for periodically sending said first updated list of addresses from a plurality of user computers to a master database situated at a location remote from the user's computer, thereby creating an updated master database of addresses received from the user computers; and

a device provided at said location remote from the user's computers for including addresses of said updated master database only if the same address is sent to said master database by a predetermined number of user computers.

8. A system for filtering electronic mail transmitted on a communication medium comprising:

one or more user computers connected to the communications medium, each of said computers including a first filter provided in each of said computers provided with a first list of addresses from which a user does not wish to receive electronic mail, a second filter provided in said computers provided with a second list of addresses from which the user does wish to receive electronic mail, a comparison device provided in each of the computers for comprising the addresses included in said first filter and said second filter with the address of electronic mail received by said computer, a first folder for retaining electronic mail received in said computer whose address is not included in said first list of addresses but is included in said second list of addresses, a second folder for retaining electronic mail received in said computer whose address is not included in either said first list of addresses or said second list of addresses for review by the user at a later time, and further including an updated device for adding addresses to said first filter based upon the user's review, thereby creating a first updated first list of addresses; and

master server and master database connected to said one or more computers by the communication medium for periodically receiving said first updated first list of addresses from each of said user computers to be inserted into said master database thereby creating an updated master database of addresses received from said user computer and for periodically transmitting to each of said user computers a list of addresses to be added to said first filter to create second updated list of first addresses; and

a device provided at said location remote from the user's computers for including addresses of said updated master database only if the same address is sent to said master database by a predetermined number of user computers.

9. A method for filtering electronic mail received over a communications medium at a user's computer comprising:

creating a first list of addresses from which each user does not wish to receive electronic mail;

inputting said first list of addresses into the user's computer as a first filter;

creating a second list of addresses from which the user does wish to receive electronic mail

inputting said second list of addresses into the user's computer as a second filter;

comparing the addresses included in said first and second list of addresses with the address of a received electronic mail;

automatically deleting the electronic mail form the user's computer if the address of the electronic mail is included in said first list of addresses;

retaining the electronic mail in a first location in the computer if the address of the electronic mail is in said second list of addresses or retaining the electronic mail in a second location in the computer if the address of the electronic mail is not included in either said first list of addresses or said second list of addresses;

reviewing any electronic mail included in said second location;

eliminating any unwanted electronic mail based upon said reviewing step;

recording the electronic address of any unwanted electronic mail in said first filter to create a first updated list of first addresses;

periodically transmitting said first updated list of first addresses to a master

database situated at a remote location to create an updated master database of addresses received from a plurality of user computers; and

wherein any address in said first updated list of first addresses is added to said master database to create said undated master database only if the said address is sent to said master database by a predetermined number of user commuters.

11. The method in accordance with claim 9, wherein said first filter includes a list of objectionable character strings compared to character strings included in a received electronic mail.

12. A system for filtering electronic mail for a plurality of user computers received over a communication medium to each of the user computers comprising:

a first filter provided in each of the user computer provided with first a list of objectionable character strings from which the user does not wish to receive electronic mail if at least one of said objectionable character strings is included in the electronic mail;

a second filter provided in each of the user computers provided with a second list of acceptable character strings from which the user does wish to receive electronic mail if at least one of said acceptable character strings is included in the electronic mail;

comparison device provided in each of the user computers for comparing the character strings included in said first filter and said second filter with the subject line and body of electronic mail received by the user computers, wherein any electronic mail which includes a character string discovered by said comparison device to be included in said first list of objectionable character strings will be automatically deleted from that particular user computer, and further wherein any electronic mail which includes a character string discovered by said comparison device to be included in said second list of character strings would be retained in the user's computer for review by the user;

a first folder for retaining electronic mail received in the computer not included in character strings provided in either said first or second lists for review by the user at a later time, and further including a update device for adding character strings to said first filter or said second filter based upon the user's review, thereby creating a first updated first list of objectionable character strings and an updated second list of acceptable character strings;

a device for periodically sending said first updated first list of objectionable character strings to a master database situated at a location remote from the user's computer thereby creating an updated master database of objectionable character strings received from the user computers; and

a device provided at said location remote from the user's computer for including objectionable character strings in said updated master database only if the same objectionable character string is sent to said master database by a predetermined number of user computers.

15. A system for filtering electronic mail for a plurality of user computers received over a communication medium by each of the user computers comprising:

a first filter located at a location remote from the user computers provided with first a list of addresses from which the user does not wish to receive electronic mail;

a second filter located at the location provided remote from the user computers with a second list of addresses from which the user does wish to receive electronic mail;

comparison device located at the location remote from the user computers for comparing the addresses included in said first filter and said second filter with the address of electronic mail received at the location remote from the user

computers and directed to the user, wherein any electronic mail whose address is discovered by said comparison device to be included in said first list of addresses will be automatically deleted from location remote from the user computers, and further wherein any electronic mail whose address is discovered by said comparison device to be included in said second list of addresses would be transmitted from the location remote from the user computers to a particular user computer for review by the user;

a first folder for retaining electronic mail received at the location remote from the user computers whose address is not included in either said first list of addresses or said second list of addresses for review by the user at a later time, and further including an update device for adding addresses to said first filter or said second filter based upon the user's review, thereby creating a first updated list of first addresses and a first updated list of second addresses;

a device for periodically sending said first updated list of first addresses from a plurality of user computers to a master database situated at a location remote from the user computers thereby creating an updated master database of addresses produced by the user computers; and

a device provided at said location remote from the user's computers for including addresses of said updated master database only if the same address is sent to said master database by a predetermined number of user computers.

18. A system for filtering electronic mail remote from the user computers received over a communication medium by each of the user computers comprising:

a first filter located at a location remote from the user computers provided with a first list of addresses from which the user does not wish to receive electronic mail as well as a first list of objectionable character strings;

a second filter located at the location remote from the user computers provided with a second list of addresses from which the user does wish to receive electronic mail as well as a second list of acceptable character strings;

comparison device located at the location remote from the user computers for comparing the addresses and character strings included in said first filter and said second filter with the address of electronic mail as well as character strings provided in the electronic mail received at the location remote from the users computers, wherein any electronic mail whose address is discovered by said comparison device to be included in said first list of addresses or electronic mail containing a character string included in the first list of objectionable character strings, will be automatically deleted from the location remote from the user computers, and further wherein any electronic mail whose address is discovered by said comparison device to be included on said second list of addresses or said second list of acceptable character strings, would be transmitted from the location remote from the user computers to a particular user computer for review by the user;

a first folder for retaining electronic mail received at the location remote from the user computers whose address is not included in said first list of addresses, said first list of objectionable character strings, said second list of addresses or said second list of acceptable character strings, for review by the user at a later time, and further including an update device for adding addresses to said first filter or said second filter based upon the user's review, thereby creating a first updated list of addresses and a first updated list of second addresses;

a device for periodically sending said first updated list of addresses from a plurality of user computers to a master database situated at a location remote from the user computers thereby creating an updated master database having addresses and character strings provided by the user computers; and

a device provided at said location remote from the user's computers for including addresses of said updated master database only if the same address is sent to said master database by a predetermined number of user computers.

21. A method for filtering electronic mail received over a communications medium be reviewed by a user computer comprising, in a system including a plurality of users computers:

creating a first list of addresses at a location removed from the user computers and connected to the communications medium thereby creating a first filter;

creating a second list of addresses from which the user does wish to receive electronic mail;

introducing said second list of addresses to the location removed from the user computers connected to the communication medium, thereby creating a second filter;

comparing the addresses included in said first and second list of addresses with the address of an electronic mail received at the location removed from the user computers;

automatically deleting the electronic mail from the location removed from the user computers if the address of the electronic mail is included in said first list of addresses;

retaining the electronic mail in a first location in the location removed from the user computers if the address of the electronic mail is in said second list of addresses or retaining the electronic mail in a second location in the location removed from the user computers if the address of the electronic mail is not included in either said first list of addresses or said second list of addresses;

reviewing any electronic mail included in said second location;

eliminating any unwanted electronic mail based upon said reviewing step;

recording the electronic address of any unwanted electronic mail in said first filter to create a first updated list of first addresses;

periodically transmitting said first updated list of first addresses to a master database situated at a remote location to create an updated master database of addresses received from a plurality of user computers; and

wherein any address in said first updated list of first addresses is added to said master database to create said updated master database only if the said address is sent to said master database by a predetermined number of user computers.

23. The method in accordance with claim 21, wherein said first filter includes a list of objectionable character strings compared to character strings included in a received electronic mail.

☐ 10. Document ID: US 5999932 A

L5: Entry 10 of 12 File: USPT Dec 7, 1999

DOCUMENT-IDENTIFIER: US 5999932 A
TITLE: System and method for filtering unsolicited electronic mail messages using data matching and heuristic processing

Abstract Text (1):
A system for eliminating unsolicited electronic mail generates and stores a user

inclusion list including identification data for identifying e-mail desired by the user. Data from one or more fields of incoming electronic mail messages are compared with the identification data stored in the user inclusion list. If the electronic mail message data matches corresponding identification data from the user inclusion list, the e-mail message is marked with a first display code, such as "OK." If no match is detected, the system performs at least one heuristic process to determine whether the electronic mail message may be of interest to the user. If the message satisfies one or more criteria as determined by the heuristic process and is therefore of potential interest to the user, the message is marked with a second display code, such as "NEW." If the e-mail message does not satisfy any of the heuristic criteria, the e-mail message may be marked with a third display code, such as "JUNK." The processed e-mail messages are displayed to the user in a display mode corresponding to the display codes respectively assigned to the messages.

Brief Summary Text (2):
The present invention relates to a method and system for filtering electronic mail ("e-mail") sent to one or more users via a communications network to eliminate unsolicited e-mail from the user's electronic mailbox. The method and system according to the present invention sort e-mail messages by comparing one or more predetermined data fields of each e-mail message with data stored in an automatically updated database of acceptable addresses and domains. The e-mail messages with matching data are forwarded to the respective user's mailbox. The e-mail messages without matching data are sorted using one or more heuristic sorting methods and categorized either as "junk," which are not of interest to the user, or as "new," which are of potential interest to the user. Each message is displayed to the user in accordance with its respective status.

Brief Summary Text (13):
(c) comparing data from the received electronic mail message with identification data in the user inclusion list;

Brief Summary Text (14):
(d) upon identifying a match between the electronic mail message data and the identification data, marking the e-mail message with a first display code; and

Brief Summary Text (17):
(f) upon failing to detect a match between the electronic mail message data and the identification data in the user inclusion list, performing at least one heuristic process to determine whether the electronic mail message may be of interest to the user;

Detailed Description Text (5):
According to one embodiment of the present invention, the e-mail filter 104 filters incoming mail received in the user's e-mail store 106 based upon three fields of data contained in the incoming e-mail, the "FROM" field, the "TO" field and the "SUBJECT" field. Notably, filtering may also include the "CC" field and the "BCC" field to filter e-mail messages on which the user is listed as a CC or BCC recipient rather than a direct recipient. Preferably, the e-mail filter 104 compares the "FROM", "TO", "CC", "BCC", and "SUBJECT" fields of an incoming e-mail message with the corresponding data categories stored in the inclusion list manager 102.

Detailed Description Text (6):
In a preferred embodiment, if the data in any of the fields of the incoming e-mail message match data in the corresponding data category stored in the inclusion list manager 102, the e-mail is marked by the filter 104 with a first display code indicating the "OK" status of the message. The marking of the incoming e-mail may be accomplished using known programming techniques as would be know to one of skill in the art, for example, by adding an additional field of information to the received e-mail format or by altering one or more existing e-mail fields to indicate the display status of the e-mail. The e-mail message is then displayed in the user's inbox by the user interface 108 in accordance with the first display code.

Detailed Description Text (7):
If the e-mail filter 104 does not detect a match between the stored inclusion list data and the data from the received e-mail message, the incoming e-mail is further

processed using one or more heuristic processing techniques to determine whether the e-mail may be of interest to the user. The filtering process and the heuristic processing techniques will be described in further detail below. If the e-mail message satisfies one or more criteria as determined with the heuristic processing, the e-mail message is marked with a second display code. If the data in the e-mail message do not match the data in the inclusion list and if the message also does not satisfy the heuristic processing criteria, then the message is marked with a third display code.

Detailed Description Text (13):
In a preferred embodiment, the list processor 201 also automatically updates the inclusion list. In order to insure that the inclusion list remains current, the list processor 201 accesses (polls) e-mail address information from the sources 203 to 208 at predetermined intervals of time such as hourly, daily, weekly or monthly. The update process may also be implemented as an interrupt-driven process prompted by one or more of the sources 203 through 208. The list processor 201 compares the e-mail addresses stored in sources 203 to 208 with those stored in the user inclusion list and adds new e-mail addresses from the sources 203 to 208 to the inclusion list. In this way, the user inclusion list may be automatically updated.

Detailed Description Text (17):
According to a preferred embodiment of the present invention, the "SUBJECT" category of the user inclusion list may be initially set to automatically include the information in the "SUBJECT" field of each message in the user's e-mail outbox. Thus, any incoming e-mail messages having "SUBJECT" field data matching the "SUBJECT" data of a message in the user's outbox, for example, a reply message, would be displayed in the user's inbox.

Detailed Description Text (19):
The "SUBJECT" data stored in the user inclusion list may be compared to the "SUBJECT" field of the incoming e-mail message, for example, using a text search, keyword search or other search as would be apparent to one of skill in the art. As with the "FROM" category of the inclusion list, the user may manually modify the "SUBJECT" category to add or delete subjects, words or phrases as desired. This category of the inclusion list may also be automatically updated to include new subject data of newly sent e-mail messages in the user's outbox. Other sources of "SUJBJECT" data, such as the user's inbox or data stored in other software programs on the user's computer may also be used to create and maintain the "SUBJECT" category of the user inclusion list stored by inclusion list processor 102.

Detailed Description Text (20):
In addition to the automatic and manual updating of the user inclusion list described above, new data may optionally be added to the user inclusion list as incoming e-mail messages are processed. For example, if a received e-mail message has "SUBJECT" or other user-definable header field data matching "SUBJECT" or other user-definable header data in the inclusion list, the "FROM" and "TO" data from the e-mail message may be automatically added to the user inclusion list by the list processor 201. As another example, when a received e-mail message has "TO" field data matching "TO" data in the inclusion list, the "FROM" and "SUBJECT" [or subset of "SUBJECT"] or user definable header data from the e-mail message may be automatically added to the inclusion list. As a further example, when a received e-mail message has "FROM" field data matching "FROM" data in the inclusion list, the "TO" and "SUBJECT" or user-definable header data may be automatically added to the inclusion list. In these and other manners apparent to users and others of ordinary skill, the inclusion list may be continually and dynamically varied as e-mail messages are received and processed.

Detailed Description Text (25):
In the preferred embodiment of the present invention depicted in FIG. 3, the filtering process performed for each user A, B, C, and D by the e-mail filter 304 is the same as that performed by filter 104 in FIG. 1. The filter 304 compares the data stored in the "TO," "FROM," "CC," "BCC," and "SUBJECT" fields of the incoming e-mail messages with corresponding categories of data stored in the inclusion list processor 302. If data in any of these fields of the incoming e-mail matches data stored in a corresponding field of the inclusion list processor 302, the incoming

e-mail is marked "OK" and forwarded to the user. If no match is detected, the e-mail filter 304 performs at least one type of heuristic processing to determine whether the e-mail may be of interest to the user, and, if not, labels the e-mail message accordingly, for example, as "JUNK."

Detailed Description Text (31):
FIG. 4 provides a process flowchart illustrating the filtering steps performed by e-mail filters 104 and 304. First, in step 401, an e-mail message is received from the network either by a user site system such as the system described in FIG. 1 or by an e-mail server such as the system described in FIG. 3. Upon receipt of an e-mail message, the e-mail filter (e.g., 104 or 304) retrieves data from selected fields of the received e-mail message as shown in step 402. In step 403, the e-mail filter compares the field data retrieved from the received message with data stored in the corresponding category of the user inclusion list. In step 410, if the field data from the received message matches a data entry stored in the corresponding category of the inclusion list, the received message is marked with a first display code indicating that the status of the message is "OK". In step 411, the field data from the received message may optionally be added to the corresponding categories of data in the user inclusion list.

Detailed Description Text (32):
As shown in FIG. 4A, comparing step 403 may include a comparison of data retrieved from the "TO," "FROM" and "SUBJECT" fields of the received message. As shown in step 404, if the "FROM" field data from the received e-mail message does not match any data entry in the "FROM" category of the stored inclusion list, the "TO," "CC," and "BCC" field data from the received message is compared to the corresponding categories of data stored in the user inclusion list in step 405.

Detailed Description Text (33):
As illustrated in step 406 of FIG. 4A, if the "TO" field data from the received e-mail message does not match any data entry in the "TO" category of the stored inclusion list, the text stored in the "SUBJECT" field data of the received message is compared to the corresponding category of text data stored in the user inclusion list. If a match is found, the received message is marked with the first display code indicating that the status of the message is "OK" (step 410). The "FROM" and "TO" data from the received message may optionally be added to the corresponding categories of data in the user inclusion list (step 411).

Detailed Description Text (34):
If no matches of the "FROM," "TO," "CC," "BCC," or "SUBJECT" field data are identified in step 403 of FIG. 4 or steps 404 to 406 of FIG. 4A, in step 412 the e-mail filter performs one or more heuristic processes to determine whether the received e-mail message meets certain criteria suggesting that the message may be of interest to the user. If the e-mail message meets one or more of the heuristic criteria, in step 413 the e-mail is marked with a second display code indicating that the status of the message is "NEW." The "1TO,""FROM" and "SUBJECT" field data from the e-mail message may optionally be added to the user inclusion list by the inclusion list processor (e.g., 102 or 302) as show in step 414.

Detailed Description Text (38):
A preferred embodiment of the heuristic processing described in step 412 of FIG. 4 will now be described in additional detail. Heuristic processing according to the present invention involves evaluating the message with one or more of the following rules. The "FROM" field matches a "TO" entry in the user inclusion list;

Detailed Description Text (39):
1. The "FROM" field has a domain that matches an Internet domain of one or more entries in the "FROM" category of the user inclusion list;

Detailed Description Text (40):
2. The "FROM" field has a domain that matches one of a pre-defined list of domains that are assured to be junk-free such as corporations or government organizations.

Detailed Description Text (41):
3. The "FROM" field has a domain that matches one of a multiplicity of domains that

are input by the user.

Detailed Description Text (43):
An alternative embodiment to the heuristics includes a user-selectable option to use any of these rules. Another alternative embodiment reduces or adds these rules to either reduced the complexity of implementation or improve the quality of the filtering. Other heuristic filtering rules may also be defined to assist the e-mail filter in identifying e-mails that do not match the stored categories of the user inclusion list but are nonetheless of interest to the user.

Detailed Description Text (44):
The filtering method according to the present invention may also be implemented in combination with one or more known exclusion-based filtering methods. A preferred embodiment of such a combination method is illustrated in FIG. 6 and includes an additional filtering step 650 in which selected data fields of the received e-mail message are compared to corresponding categories in a stored exclusion list (for example, stored in inclusion list processor 104 or 304). In the preferred embodiment, if any matches are detected, the e-mail message is automatically marked "JUNK." The remainder of the method steps shown in FIG. 6. correspond to similarly numbered steps in FIG. 4.

Current US Cross Reference Classification (2):
709/206

Current US Cross Reference Classification (3):
709/207

Current US Cross Reference Classification (4):
709/315

CLAIMS:

1. A method for filtering electronic mail addressed to a user, comprising the steps of:

storing a user inclusion list including identification data for identifying e-mail desired by the user;

receiving an electronic mail message;

comparing data from said received electronic mail message with said identification data;

upon identifying a match between said electronic mail message data and said identification data, marking said electronic mail with a first display code;

displaying in a first display format said electronic mail message marked with the first display code to the user;

upon failing to detect a match between said electronic mail message data and said identification data, performing at least one heuristic process to determine whether said electronic mail message may be of interest to the user;

upon identifying an electronic mail message of interest to the user, marking said electronic mail with a second display code;

displaying said electronic mail message marked with said second display code to the user in a second display format;

upon failing to identify an electronic mail message of interest to the user, marking the electronic mail message with a third display code; and

displaying said electronic mail message marked with said third display code to the user in a third display format.

10. A method according to claim 1, wherein said at least one heuristic process includes at least one of the following tests:

(a) a test to determine whether a first field of said received electronic mail message matches a corresponding entry in said user inclusion list;

(b) a test to determine whether said first field of said received electronic mail message has a domain that matches an Internet domain of one or more entries in the corresponding category of said user inclusion list;

(c) a test to determine whether the first field of said received electronic mail message has a domain that matches one of a pre-defined list of domains; or

(d) a test to determine whether a second field of said received electronic mail message matches a second entry in said user inclusion list.

12. A method according to claim 11, wherein, when the data from said electronic mail message matches data stored in said exclusion list, said electronic mail message is marked with said third display code and displayed to the user in said third display mode.

14. A system for eliminating unsolicited electronic mail, comprising:

an inclusion list processor for storing identification data for identifying e-mail desired by the user;

an e-mail storage unit for storing incoming electronic mail messages;

an e-mail filter for filtering said stored incoming electronic mail messages in accordance with said identification data stored in said inclusion list processor and for marking each of said electronic mail messages with one of a plurality of display codes to indicate a status of each of said messages; and

a user interface for displaying said filtered electronic mail messages to a user in accordance with said display codes;

wherein said filtering performed by said e-mail filter includes the steps of

(a) comparing data from said incoming electronic mail messages with said identification data;

(b) upon identifying a match between said electronic mail message data and said identification data, marking said electronic mail with a first display code;

(c) transmitting said electronic mail message marked with the first display code to said user interface;

(d) upon failing to detect a match between said electronic mail message data and said identification data, performing at least one heuristic process to determine whether said electronic mail message may be of interest to the user;

(e) upon identifying an electronic mail message of interest to the user, marking said electronic mail message with a second display code; and

(f) transmitting said electronic mail message marked with said second display code to said user interface.

27. A system according to claim 14, wherein said at least one heuristic process includes one or more of the following tests:

(a) a test to determine whether a first field of said received electronic mail message matches a corresponding entry in said user inclusion list;

(b) a test to determine whether said first field of said received electronic mail message has a domain that matches an Internet domain of one or more entries in the

corresponding category of said user inclusion list; or

(c) a test to determine whether said first field of said received electronic mail message has a domain that matches one of a pre-defined list of domains.

28. A system according to claim 14, further comprising the step of filtering said electronic mail message using an exclusion list, wherein, when the data from said electronic mail message matches data stored in said exclusion list, said electronic mail message is marked with said third display code and displayed to the user in said third display mode.

30. A method for filtering electronic mail addressed to a user, comprising the steps of:

storing a user inclusion list including identification data for identifying e-mail desired by the user;

receiving an electronic mail message;

comparing data from said received electronic mail message with said identification data;

upon identifying a match between said electronic mail message data and said identification data, marking said electronic mail with a first display code;

displaying in a first display format said electronic mail message marked with the first display code to the user;

upon failing to detect a match between said electronic mail message data and said identification data, performing at least one heuristic process to determine whether said electronic mail message may be of interest to the user;

upon identifying an electronic mail message of interest to the user, marking said electronic mail with a second display code;

displaying said electronic mail message marked with said second display code to the user in a second display format; and

upon failing to identify an electronic mail message of interest to the user, marking the electronic mail message with a third display code such that said electronic mail message is not displayed to the user.

31. A method according to claim 30, wherein said at least one heuristic process includes at least one of the following tests:

(a) a test to determine whether a first field of said received electronic mail message matches a corresponding entry in said user inclusion list;

(b) a test to determine whether said first field of said received electronic mail message has a domain that matches an Internet domain of one or more entries in the corresponding category of said user inclusion list;

(c) a test to determine whether the first field of said received electronic mail message has a domain that matches one of a pre-defined list of domains; or

(d) a test to determine whether a second field of said received electronic mail message matches a second entry in said user inclusion list.

WEST

Generate Collection | Print

Search Results - Record(s) 11 through 12 of 12 returned.

☐ 11. Document ID: US 5978837 A

L5: Entry 11 of 12 File: USPT Nov 2, 1999

DOCUMENT-IDENTIFIER: US 5978837 A
TITLE: Intelligent pager for remotely managing E-Mail messages

Abstract Text (1):
An intelligent pager remotely manages an E-Mail server that receives E-Mail messages transmitted over an E-Mail network. Useful E-Mail messages can be identified by the pager and separated from the junk mail. The pager remotely controls the server to forward messages as directed by the recipient. The E-Mail server sends a summary of the E-Mail message to the pager over a different network so that the pager can select a forwarding destination for the message. The pager sends a selection signal to the E-mail server which then forwards the E-Mail message to the selected destination. The E-Mail message can be forwarded by the server over another network, such as the public telephone network, to a computer or a FAX machine at the recipient's office, home or other destination.

Brief Summary Text (9):
In accordance with the invention, the recipient's computer automatically makes some of the E-Mail forwarding selections for the recipient. The recipient's computer includes a local database of sender records for those senders whose messages are considered important to the recipient. A control program in the recipient's computer automatically reviews the title or message summary in an alert signal from the E-Mail server and automatically sends to the E-Mail server a selection signal designating a selected forwarding destination to send the E-Mail message. Each sender record includes a priority word list of words that the recipient considers important in any messages sent by that particular sender. The recipient's computer includes a priority word comparison program to compare the words received in the alert signal with the words in the priority word list in the sender's record. If there is no match with a priority word, then the recipient's computer sends a selection signal to the server with a pre-selected default choice for the forwarding destination. However, if there is a match with a priority word, then the recipient's computer selects a first destination choice from the sender's record to send in the selection signal to the E-Mail server.

Brief Summary Text (10):
In a further embodiment of the invention, if there is a match and an additional alarm value is present in the sender's record, then the computer uses its pager card to send an alert signal to the recipient's mobile pager. In this manner, the recipient can manually make the forwarding selection for particularly important E-Mail messages.

Brief Summary Text (11):
In an alternate embodiment of the invention, the recipient's mobile pager or personal communications device carried by the recipient automatically makes some of the E-Mail forwarding selections for the recipient. The recipient's mobile pager or personal communications device includes a local database of the sender records described above for those senders whose messages are considered important to the recipient. A control program in the recipient's mobile pager or personal communications device automatically reviews the title or message summary in an alert signal from the E-Mail server and automatically sends to the E-Mail server a

selection signal designating a selected forwarding destination to send the E-Mail message. The recipient's mobile pager or personal communications device includes a priority word comparison program to compare the words received in the alert signal with the words in the priority word list in the sender's record. If there is no match with a priority word, then the recipient's mobile pager or personal communications device sends a selection signal to the server with a pre-selected default choice for the forwarding destination. However, if there is a match with a priority word, then the recipient's mobile pager or personal communications device selects a first destination choice from the sender's record to send in the selection signal to the E-Mail server. If there is a match and an additional alarm value is present in the sender's record, then the recipient's mobile pager or personal communications device sounds an audible alarm to alert the recipient. The recipient can then manually make the forwarding selection for particularly important E-Mail messages.

Detailed Description Text (7):
The Computer 70 is shown in greater detail in FIG. 3. Memory 702 is connected by the bus 704 to the wireless pager transceiver card 71, the display and keyboard 708, the telephone modem 710, the processor 712, and the database storage 762. The memory 702 includes an operating system program 720, a telephone network communications program 722, a paging network communications program 724, a priority word comparison program 726, and the control program 728 shown in FIG. 4. The programs are composed of executable instructions which, when executed by the processor 712, carry out the functions performed by the Computer 70. Memory 702 also includes a sender record buffer 750 which stores the sender record 755 accessed from the database storage 762. The record 755 is shown with the fields for sender ID 751, the priority word list 752, the choice.sub.-- 1 address 753, the choices.sub.-- 2 address 754, and the alarm 756. Memory 702 also includes an E-Mail receive buffer 730 which stores an E-Mail message 239 received from the E-Mail server 60. The message is shown with the fields for destination address 232, sender ID 233, title 235, message data 237, and E-Mail ID 231. Memory 702 also includes an alert signal buffer 740 with the alert signal 245' having fields 233, 235, and 231. Memory 702 also includes a selection signal buffer 746 with a selection signal 249' having fields 231 and 248. Memory 702 also includes a pre-selected value for choice (no alarm) 764. Memory 702 also includes a buffer 766 to store the selected value for choice in response to page (alarm).

Detailed Description Text (8):
In accordance with the invention, the Computer 70 is controlled by the control program 728 shown in FIG. 4. Step 802 stores the pre-selected value for choice 764 and the sender records 755. Each sender record 755 includes the sender's ID, a priority word list 752 of words that the recipient considers important in any messages sent by that particular sender, a list of choices 753 and 754 for forwarding addresses, and an alarm value 756 for each sender record. Step 804 waits to receive an alert signal 245' at pager card 71. When an alert signal is received, Step 805 accesses the sender's record 755 from the database storage 762. The recipient's computer 70 includes the priority word comparison program 726 to compare the words received in the alert signal 245' with the words in the priority word list 752 in the sender's record 755. Step 806 determines if a priority word is in the alert signal. If there is no match with a priority word, then step 808 has the recipient's computer 70 send a selection signal 249' to the server 60 with a pre-selected default choice 764 for the forwarding destination. However, if Step 806 determines that there is a match with a priority word, then the program passes to step 809. Step 809 determines if the alarm value 756 is off and if it is off, then the program uses choice.sub.-- 1 753 and goes to step 814. The recipient's computer 70 selects a first destination choice 753 from the sender's record 755 to send in the selection signal 249' to the E-Mail server 60.

Detailed Description Text (9):
If there is a match and Step 809 determines that the alarm value 756 is on, then the program passes to step 810. In Step 810, since an alarm value 756 is present in the sender's record 755, the computer 70 uses its wireless pager transceiver card 71 to send an alert signal 245' to the recipient's mobile pager 74. In this manner, the recipient can manually make the forwarding selection for particularly important E-Mail messages. Step 812 receives the selected choice value 766 from the

recipient's mobile pager. Then step 814 sends the selection signal 249' with the choice value to the E-Mail server 60.

Detailed Description Text (11):
The mobile pager 74 is shown in greater detail in FIG. 5. Memory 902 is connected by the bus 904 to the radio interface 906, the display and keyboard 908 with buttons 1, 2, and 3, the audio output 910, and the processor 912. The memory 902 includes an operating system program 920, a paging network communications program 924, a priority word comparison program 926, and the control program 928 shown in FIG. 10. The programs are composed of executable instructions which, when executed by the processor 912, carry out the functions performed by the mobile pager 74. Memory 902 also includes a sender record buffer 950 which stores all of the sender records 755. A record 755 is shown with the fields for sender ID 751, the priority word list 752, the choice.sub.-- 1 address 753, the choice.sub.-- 2 address 754, and the alarm 756. Memory 902 also includes an E-Mail receive buffer 930 which stores an E-Mail message 239 received from the E-Mail server 60. The message is shown with the fields for destination address 232, sender ID 233, title 235, message data 237, and E-Mail ID 231. Memory 902 also includes an alert signal buffer 940 with the alert signal 245 having fields 233, 235, and 231. Memory 902 also includes a selection signal buffer 946 with a selection signal 249 having fields 231 and 248. Memory 902 also includes a pre-selected value for choice (no alarm) 764. Memory 902 also includes a buffer 766' to store the selected value for choice in response to pressing button 1, 2, or 3 (alarm).

Detailed Description Text (12):
In accordance with the invention, the mobile pager 74 is controlled by the control program 928 shown in FIG. 6. Step 1002 stores the pre-selected value for choice 764 and the sender records 755 in memory 902. Each sender record 755 includes the sender's ID, a priority word list 752 of words that the recipient considers important in any messages sent by that particular sender, a list of choices 753 and 754 for forwarding addresses, and an alarm value 756 for each sender record. Step 1004 waits to receive an alert signal 245 at radio interface 906. When an alert signal is received, Step 1005 accesses the sender's record 755 from the memory 902. The recipient's mobile pager 74 includes the priority word comparison program 926 to compare the words received in the alert signal 245 with the words in the priority word list 752 in the sender's record 755. Step 1006 determines if a priority word is in the alert signal. If there is no match with a priority word, then step 1008 has the recipient's mobile pager 74 send a selection signal 249 to the server 60 with a pre-selected default choice 764 for the forwarding destination. However, if Step 1006 determines that there is a match with a priority word, then the program passes to step 1009. Step 1009 determines if the alarm value 756 is off and if it is off, then the program uses choice.sub.-- 1 753 and goes to step 1014. The recipient's mobile pager 74 selects a first destination choice 753 from the sender's record 755 to send in the selection signal 249 to the E-Mail server 60.

Detailed Description Text (13):
If there is a match and Step 1009 determines that the alarm value 756 is on, then the program passes to step 1010. In Step 1010, since an alarm value 756 is present in the sender's record 755, the mobile pager 74 outputs an audio signal on the audio output 910 to alert the recipient. In this manner, the recipient can manually make the forwarding selection for particularly important E-Mail messages. Step 1012 receives the selected choice value 766' from the recipient's pressing the buttons 1, 2, or 3 on the keyboard 908. Then Step 1014 sends the selection signal 249 with the choice value to the E-Mail server 60. If the recipient's choice is to view the entire E-Mail message on the display of the mobile pager 74 or personal communications device, the server 60 will transmit the entire E-Mail message over the radio link 75 to the mobile pager 74.

Current US Original Classification (1):
709/207

CLAIMS:

1. A communications device for transmitting a selection value to an E-Mail server buffering an E-mail message received from an E-Mail network, to forward the E-Mail

message to a selected forwarding destination, comprising:

a paging system coupled directly to the E-Mail server;

a personal computer that includes a wireless pager card coupled to the paging system via a radio link;

a data storage within said personal computer for storing a pre-selected default forwarding destination and a plurality of sender records, each record being for an E-Mail sender and including the E-Mail sender's identity, a priority word list of words, an alarm value that is either on or off, and a list of forwarding destinations, which includes a first forwarding destination, for an E-Mail message from the E-Mail server;

said wireless pager card receiving an alert signal from said paging system over said radio link, the alert signal including a received identity of a sender and a title and message summary including words;

a processor within said personal computer coupled to the storage and to the wireless pager card, for comparing the received identity of the sender with the sender identities in the plurality of sender records and comparing the words in the title and message summary with the words in the priority word list;

if there is no match between the words in the title and message summary and the words in the priority word list, said processor outputting said pre-selected default forwarding destination as said selected forwarding destination for the E-Mail message;

if there is a match between the words in the title and message summary and the words in the priority word list, and if said alarm value is off, said processor outputting said first forwarding destination in a sender record having a sender identity that matches the received sender identity as said selected forwarding destination for the E-Mail message;

if there is a match between the words in the title and message summary and the words in the priority word list, and if said alarm value is on, said processor outputting said list of forwarding destinations in a sender record having a sender identity that matches the received sender identity;

a display on said personal computer displaying said list of forwarding destinations;

a key on said personal computer whereby one of said list of forwarding destinations can be selected by depressing said key;

said wireless pager card sending a selection signal to said paging system over said radio link, wherein said selection signal includes a selection value that indicates said selected forwarding destination for the E-Mail message; and

wherein said selection value is transferred to said E-Mail server.

2. In a communications device for transmitting a selection value to an E-Mail server buffering an E-Mail message received from an E-Mail network, to forward the E-Mail message to a selected forwarding destination, a method comprising the steps of:

storing a pre-selected default forwarding destination and a plurality of sender records in a pager, each record being for an E-Mail sender and including the E-Mail sender's identity, a priority word list of words, an alarm value that is either on or off, and a list of forwarding destinations, which includes a first forwarding destination, for an E-Mail message from the E-Mail server;

receiving an alert signal over a radio link from a paging system that is coupled directly to the E-Mail server, the alert signal including a received identity of a sender and a title and message summary including words;

comparing in the pager the received identity of the sender with the sender identities in the plurality of sender records and comparing in the pager the words in the title and message summary with the words in the priority word list;

if there is no match between the words in the title and message summary and the words in the priority word list, outputting said pre-selected default forwarding destination as said selected forwarding destination for the E-Mail message;

if there is a match between the words in the title and message summary and the words in the priority word list, and if said alarm value is off, outputting said first forwarding destination in a sender record having a sender identity that matches the received sender identity as said selected forwarding destination for the E-Mail message;

if there is a match between the words in the title and message summary and the words in the priority word list, and if said alarm value is on, (i) outputting said list of forwarding destinations in a sender record having a sender identity that matches the received sender identity, (ii) displaying said list of forwarding destinations on said pager, and (iii) selecting one of said list of forwarding destinations by depressing a button on said pager;

sending a selection signal to said pager system over said radio link, wherein said selection signal includes a selection value that indicates said selected forwarding destination for the E-Mail message; and

transferring said selection value to said E-Mail server.

3. An article of manufacture for use in a computer system, comprising:

a computer useable medium having computer readable program code means embodied therein for transmitting a selection value to an E-Mail server buffering an E-Mail message received from an E-Mail network, to forward the E-mail message to a selected forwarding destination, the computer readable program code means in said article of manufacture comprising:

computer readable program code means for causing a computer to store a pre-selected default forwarding destination and a plurality of sender records, each record being for an E-Mail sender and including the E-Mail sender's identity, a priority word list of words, an alarm value that is either on or off, and a list of forwarding destinations, which includes a first forwarding destination, for an E-Mail message from the E-Mail server;

computer readable program code means for causing a computer to receive an alert signal over a radio link from a paging system coupled directly to the E-Mail server, the alert signal including a received identity of a sender and a title and message summary including words;

computer readable program code means for causing a computer to compare the received identity of the sender with the sender identities in the plurality of sender records and compare the words in the title and message summary with the words in the priority word list;

computer readable program code means for causing a computer to output said pre-selected default forwarding destination as said selected forwarding destination for the E-Mail message if there is no match between the words in the title and message summary and the words in the priority word list;

computer readable program code means for causing a computer to output said first forwarding destination in a sender record having a sender identity that matches the received sender identity as said selected forwarding destination for the E-Mail message if there is a match between the words in the title and message summary and the words in the priority word list, and if the alarm value is off;

computer readable program code means for causing a computer to output said list of forwarding destinations in a sender record having a sender identity that matches the

received sender identity if there is a match between the words in the title and message summary and the words in the priority word list, and if the alarm value is on;

computer readable program code means for causing a computer to display said list of forwarding destinations;

computer readable program code means for causing a computer to respond to a key on the computer being depressed to select one of said list of forwarding destinations;

computer readable program code means for causing a computer to send a selection signal over said radio link to the paging system, wherein said selection signal includes a selection value that indicates said selected forwarding destination for the E-Mail message; and

computer readable program code means for causing a computer to transfer said selection value to said E-Mail server.

4. A communications system, comprising:

a telecommunications transmission system;

an E-Mail server buffering an E-Mail message received from an E-Mail network while waiting for a recipient to select a forwarding destination in the telecommunications transmission system for the E-Mail message;

a paging system coupled directly to the E-Mail server;

a pager coupled to the paging system via a radio link;

a data storage within said pager for storing a pre-selected default forwarding destination and a plurality of sender records, each record being for an E-Mail sender and including the E-Mail sender's identity, a priority word list of words, an alarm value that is either on or off, and a list of forwarding destinations, which includes a first forwarding destination, in the telecommunications transmission system for an E-Mail message from the E-Mail server;

a receiver within said pager for receiving an alert signal over said radio link from said paging system, the alert signal including a received identity of a sender and a title and message summary including words;

a processor within said pager coupled to the storage and to the receiver, for comparing the received identity of the sender with the sender identities in the plurality of sender records and comparing the words in the title and message summary with the words in the priority list of words;

if there is no match between the words in the title and message summary and the words in the priority word list, said processor outputting said pre-selected default forwarding destination;

if there is a match between the words in the title and message summary and the words in the priority word list, and if the alarm value is off, said processor outputting said first forwarding destination in a sender record having a sender identity that matches the received sender identity;

if there is a match between the words in the title and message summary and the words in the priority word list, and if the alarm value is on, said processor outputting said list of forwarding destinations in a sender record having a sender identity that matches the received sender identity;

a display on said pager for displaying said list of forwarding destinations;

a button on said pager for selecting one of said list of forwarding destinations by depressing said button;

a transmitter within said pager coupled to the processor, for sending a selection signal to said paging system, wherein said selection signal includes a selection value that indicates said selected forwarding destination for the E-Mail message.

☐ 12. Document ID: US 5619648 A

L5: Entry 12 of 12 File: USPT Apr 8, 1997

DOCUMENT-IDENTIFIER: US 5619648 A
TITLE: Message filtering techniques

Abstract Text (1):
Techniques for reducing the amount of junk e-mail received by a user of an e-mail system. A recipient description containing non-address information is added to an e-mail message. The user has an e-mail filter which has access to information which provides a model of the user. The e-mail filter uses the non-address information and the model information to determine whether the e-mail message should be provided to the user. The e-mail filter further has access to information which provides models of the user's correspondents. If the filter does not provide the message to the user, it uses the non-address information and the model information of the user's correspondents to determine who the message might be forwarded to. A sender of e-mail can also use the model information of the sender's correspondents together with the non-address information to determine who the message should be sent to. The techniques are used in a system for locating expertise.

Brief Summary Text (12):
The first and second aspects of the invention are combined in a further aspect of the invention, namely a system for locating expertise in the e-mail system. In this system, the sender specifies an area of expertise by means of a list of keywords which are relevant to the area. The list of keywords is included in a recipient specifier in the message. The mail filter for a potential recipient has access to the document files of the potential recipient and to a list of the e-mail messages sent and received by the potential recipient. The mail filter uses the document files to determine the recipient's areas of expertise. If the keywords in the recipient specifier match one of the areas of expertise, the mail filter provides the e-mail message to the potential recipient; if not, the mail filter uses the list of e-mail messages to determine correspondents of the the potential recipient who may have the area of expertise specified in the recipient specifier and forwards the message to those correspondents. The mail filter of each potential recipient which actually provides the message to the recipient further sends a referral message to the sender of the message, who thus knows exactly who received the message.

Detailed Description Text (23):
When I have an expertise location need, I present the problem to my mail filter as an unstructured text description. Again using IR techniques, my mail filter selects a medium-to-large set of my contacts to whom the query may be relevant. It then broadcasts the query, not to the people themselves, but to their mail filters. Upon receipt of the question, each mail filter checks if its owner's user model does indeed provide a good match. If there is a good match, the mail filter presents my request to its owner. If the owner's model does not match, but the model of one of the owner's contacts does, then the mail filter can ask the owner if it can provide a referral. Finally, if there is no match at all, the query is silently logged and deleted. A great deal of flexibility can be built into each mail filter, depending upon its owner's preferences. For example, I might allow automatic referrals to be given to requests that come from my closest colleagues.

Detailed Description Text (40):

Second, names are added that result from the following computation. First, for each word that appears in the expertise request, mail filter 109 retrieves from email inverted index file 213 a list of messages 403(0 . . . n) (FIG. 4) containing that word. Next, the intersection of the lists is computed, generating a list of messages 405 each of which appears in every one of the previous lists. Next, list of messages 405 is compared against sender/recipient list file 221, and the total number of messages that appear in list of messages 405 that are from each person in sender/recipient list 221 is calculated. The result is a name/message number pair list 407 of pairs of "person name" and "number of messages". Finally, list 407 is sorted according to "number of messages". The 20 names with the highest number of messages in this list are then added to the list of candidates.

Detailed Description Text (43):
First, the words in expertise description 401 contained in the message's recipient specifier 121 are matched against the recipient's user expertise list 301. If the words appear in list 301, then mail filter 109 assumes that this request is appropriate for the recipient to see.

Detailed Description Text (44):
If the words in the phrase do not match against the contents of user expertise list 301, mail filter 109 uses user files inverted index file 305 to match the phrase against the contents of all of the recipient's files which are indexed in file 305. This matching can be efficiently performed using the program "GLIMPSE" mentioned above. If the number of matches is greater then a threshold number (e.g., more than 10 matches), the recipient's mail filter 109 determines that this request is likely to be appropriate for the recipient.

Detailed Description Text (50):
1. You wish to contact a large number of people, without necessarily broadcasting messages to everyone in the world. In the expertise location example, the user agent helped determine a preliminary list of candidates based on a matching scheme. Other ways of determining whom to send the message to are also useful. In the example below, the recipients are simply taken to be a fixed list of the sender's friends and colleagues.

Current US Original Classification (1):
709/206

CLAIMS:

4. An arrangement for locating expertise in a messaging system implemented in a computer system, comprising:

first means, included in a message, for indicating, via non-address information, expertise sought by a sender of the message;

second means in the computer system, for determining expertise of an addressee of the message;

third means in the computer system responsive to receipt of the message, for determining whether the expertise indicated by the first means matches the expertise of the addressee determined by the second means;

fourth means in the computer system responsive to a determination by the third means that the indicated expertise matches the determined expertise, for providing the message to the addressee, and responsive to a determination by the. third means that the indicated expertise does not match the determined expertise, for preventing the message from being provided to the addressee;

fifth means in the computer system, for determining expertise of contacts of the addressee;

sixth means responsive to a determination that the indicated expertise does not match the determined expertise of the addressee, for determining whether the indicated expertise matches the expertise of any said contacts determined by the

fifth means; and

seventh means responsive to a determination by the sixth means that the indicated expertise matches the determined expertise of a contact, for sending the message to that contact.

7. The arrangement of claim 4 further comprising:

eighth means responsive to a determination by the sixth means that the indicated expertise does not match the determined expertise of any contact, for discarding the message.

10. The arrangement of claim 4 further comprising:

eighth means in the computer system responsive to the sixth means determining that the indicated expertise matches the determined expertise of a contact, for including referral information in the message to indicate that the message is being sent from the addressee to that contact.

12. An arrangement for locating expertise in a messaging system implemented in a computer system, comprising:

first means, included in a message, for indicating, via non-address information, expertise sought by a sender of the message;

second means in the computer system, for analyzing files of an addressess of the message to determine therefrom expertise of the addressee;

third means in the computer system responsive to receipt of the message, for determining whether the expertise indicated by the first means matches the expertise of the addressee determined by the second means; and

fourth means in the computer system responsive to a determination by the third means that the indicated expertise matches the determined expertise, for providing the message to the addressee, and responsive to a determination by the third means that the indicated expertise does not match the determined expertise, for preventing the message from being provided to the addressee.

13. An arrangement for locating expertise in a messaging system implemented in a computer system, comprising:

first means, included in a message, for indicating, via non-address information, expertise sought by a sender of the message;

second means in the computer system, for determining expertise of an addressee of the message;

third means in the computer system responsive to receipt of the message, for determining whether the expertise indicated by the first means matches the expertise of the addressee determined by the second means;

fourth means is in the computer system responsive to a determination by the third means that the indicated expertise matches the determined expertise, for providing the message to the addressee, and responsive to a determination by the third means that the indicated expertise does not match the determined expertise, for preventing the message from being provided to the addressee;

fifth means in the computer system for analyzing messages exchanged by the sender with potential recipients of the message to determine therefrom the expertise of the potential recipients; and

sixth means in the computer system responsive to generation of the message by the sender, for selecting addressees of the message from the potential recipients by matching the expertise sought by the sender with the expertise of the potential recipients determined by the fifth means.

16. An arrangement for locating expertise in a messaging system implemented in a computer system, comprising:

first means, included in a message, for indicating, via non-address information, expertise sought by a sender of the message;

second means in the computer system, for determining expertise of an addressee of the message;

third means in the computer system responsive to receipt of the message, for determining whether the expertise indicated by the first means matches the expertise of the addressee determined by the second means;

fourth means in the computer system responsive to a determination by the third means that the indicated expertise matches the determined expertise, for providing the message to the addressee, and responsive to a determination by the third means that the indicated expertise does not match the determined expertise, for preventing the message from being provided to the addressee; and

fifth means in the computer system responsive to the fourth means providing the message to the addressee, for sending a referral message to the sender to inform the sender that the message was provided to the addressee.

Term	Documents
COMPAR$	0
COMPAR.USPT.	3498
COMPARA.USPT.	3883
COMPARAATIVE.USPT.	3
COMPARAATOR.USPT.	5
COMPARABALE.USPT.	5
COMPARABE.USPT.	7
COMPARABEL.USPT.	1
COMPARABIE.USPT.	1
COMPARABIITY.USPT.	1
COMPARABILITIES.USPT.	6
(L4 AND (COMPAR$ OR MATCH$)).USPT.	12

There are more results than shown above. Click here to view the entire set.

Display Format: KWIC | Change Format

Previous Page Next Page